

ADVERTIMENT. L'accés als continguts d'aquesta tesi queda condicionat a l'acceptació de les condicions d'ús establertes per la següent llicència Creative Commons:  <https://creativecommons.org/licenses/?lang=ca>

ADVERTENCIA. El acceso a los contenidos de esta tesis queda condicionado a la aceptación de las condiciones de uso establecidas por la siguiente licencia Creative Commons:  <https://creativecommons.org/licenses/?lang=es>

WARNING. The access to the contents of this doctoral thesis it is limited to the acceptance of the use conditions set by the following Creative Commons license:  <https://creativecommons.org/licenses/?lang=en>

Tesi Doctoral

Discriminació algorítmica: riscos, regulació, garanties jurisdiccionals i garanties
institucionals

Anna Capellà Ricart

Director i tutor: Dr. Antoni Roig Batalla

Universitat Autònoma de Barcelona

Facultat de Dret

Departament de Dret Públic i Ciències Historicojurídiques

2023

AGRAÏMENTS

Al Dr. Antoni Roig Batalla per la seva direcció i tutorització.

A en Joan Grimalt Álvaro, la Montse Ricart Martí, la Pepa Llovera Noves, en Salva Capellà Llovera, la Tatiana González Gerster i la Xon Ricart Martí, pel seu ajut i el seu suport.

INTRODUCCIÓ.....	9
PRIMERA PART. Límits de la regulació europea per protegir les persones de la discriminació causada per decisions automatitzades.....	13
Capítol I. Per què les decisions automatitzades poden ser discriminatòries?.....	16
1. Conceptes introductoris	17
1.1. La igualtat i els sistemes algorítmics.....	17
1.2. L'equitat i els sistemes algorítmics.....	20
1.3. La discriminació i els sistemes algorítmics	25
2. Causes de discriminació en les decisions automatitzades mitjançant sistemes algorítmics	29
2.1. Discriminació en base a mesures tècniques.....	31
2.2. Discriminació en base a apreciacions humanes.....	48
3. Factors que caracteritzen la discriminació algorítmica	50
3.1. Els biaixos algorítmics com a font de la discriminació	50
3.2. Opacitat intrínseca de les decisions automatitzades: una manera d'emparar la discriminació	58
3.3. Personalització en els sistemes algorítmics	61
3.4. Magnificació dels efectes de la discriminació	63
4. Conclusions parcials	65
Capítol II. Els límits de la regulació de la discriminació algorítmica en la normativa europea antidiscriminació.....	67
1. Categories protegides i àmbits d'aplicació material.....	67
1.1. Introducció al dret antidiscriminatori europeu	68
1.2. Protecció en base a certs atributs protegits.....	73
1.3. Directives antidiscriminació europees.....	77
2. Discriminació algorítmica com a conseqüència de sistemes d'opressió sobre grups de persones	83
2.1. Com posicionar la discriminació algorítmica dins la perspectiva del dret antidiscriminatori anticlassificació i antisubordinació	83
2.2. Discriminació algorítmica en base a grups estructuralment oprimits.....	86
3. Deficiències de la conceptualització bidimensional entre discriminació directa i discriminació indirecta	91
3.1. Discriminació algorítmica directa	93
3.2. Discriminació algorítmica indirecta	95
3.3. Per què la discriminació algorítmica es tradueix sobretot en discriminacions indirectes?.....	107
4. Deficiències respecte a la prova en el dret antidiscriminatori europeu per abordar la discriminació mitjançant sistemes algorítmics	115
4.1. Repartiment de la càrrega de la prova	116

4.2. Dificultat d'establir la responsabilitat de l'acció discriminatòria.....	121
4.3. Dificultat d'establir un cas <i>prima facie</i> de discriminació algorítmica d'acord amb els criteris jurisprudencials establerts pel TJUE i el TEDH	123
5. Conclusions parcials	130
Capítol III. Els límits de la regulació de la discriminació mitjançant decisions automatitzades en l'RGPD	133
1. Una protecció dèbil i dispersa: article 22 de l'RGPD	137
1.1. Decisions automatitzades i elaboració de perfils.....	138
1.2. Excepcions a la prohibició general.....	145
1.3. Establiment de mesures adequades per salvaguardar els drets, llibertats i interessos legítims de les persones interessades	146
1.4. Vaguetat de l'article 22 de l'RGPD	148
2. Biaixos algorítmics	151
2.1. Absència de referència als biaixos en l'articulat	151
2.2. L'RGPD i l'aposta de la protecció en base a categories especials de dades personals	152
3. Informació i transparència	156
3.1. Com es regula l'accés a la informació i la transparència en l'RGPD?.....	157
3.2. Obstacles per l'obtenció d'informació i transparència a causa de les decisions automatitzades mitjançant sistemes algorítmics.....	159
4. Límits a la gestió del risc de discriminació algorítmica per part de l'RGPD...	164
4.1. L'avaluació d'impacte relativa a la protecció de dades	166
4.2. Sistemes d'auditories insuficients	170
5. Perspectiva individual de la protecció que confereix l'RGPD	173
5.1. Limitacions de la perspectiva individual de la protecció	174
5.2. Drets individuals de l'article 22.3 de l'RGPD.....	176
6. Conclusions parcials	184
SEGONA PART. Potencialitats de la regulació europea per protegir a les persones de la discriminació causada per decisions automatitzades.....	187
Capítol IV. Possibilitats de la regulació antidiscriminació per abordar la discriminació algorítmica.....	188
1. Llistes obertes d'atributs protegits i homogeneïtzació de la protecció.....	188
1.2. Homogeneïtzació de la protecció	194
2. Reconeixement del caràcter grupal en la discriminació algorítmica.....	197
2.1. Reconeixement de la discriminació estructural.....	198
2.2. Reconeixement de la discriminació interseccional.....	200
3. Flexibilització dels conceptes discriminació directa i discriminació indirecta	206
3.1. Elements d'interès de cada tipologia de discriminació	206
3.2. Discriminació per associació i per error.....	215

4. Possibilitats de millora de la prova.....	225
4.1. No necessitat de condició de víctima (cas Feryn) i suficiència d'un dany hipotètic (cas Accept i Rete Lenford).....	225
4.2. Accés a la informació: Cas Meister.....	229
5. Conclusions parcials.....	232
Capítol V. Aportacions de l'RGPD a la protecció contra la discriminació algorítmica.....	236
1. Possibilitats de l'RGPD com a sistema de governança col·laborativa per abordar la discriminació algorítmica.....	237
1.1. L'RGPD com a sistema de governança col·laborativa.....	237
1.2. Condicions per l'efectivitat d'un règim de governança col·laborativa.....	239
2. Qualitat de les dades per abordar els biaixos algorítmics.....	245
2.1. L'exactitud.....	246
2.2. Drets de les persones sobre les seves dades: accés, supressió, rectificació, limitació del termini de conservació i portabilitat.....	248
2.3. Capitalitzar les dades.....	253
3. Informació útil i senzilla per fer front a la discriminació algorítmica.....	255
3.1. El dret a una explicació en la protecció envers la discriminació.....	256
3.2. Contingut de la informació individual necessària per iniciar un plet de discriminació.....	261
3.3. Política de transparència de l'RGPD.....	265
4. Responsabilitat proactiva plantejada en l'RGPD.....	268
4.1. Promoció d'una política del compliment i responsabilitat de l'acció discriminatòria.....	269
4.2. Sistemes d'autoregulació: Codis de conducta, certificacions, segells i marques.....	275
4.3. Igualtat de tractament des del disseny i per defecte.....	279
4.4. De les avaluacions d'impacte de protecció de dades a les avaluacions d'impacte algorítmiques.....	283
5. Vessant col·lectiva de la protecció envers la discriminació: Article 80 RGPD.....	293
5.1. Article 80.2 RGPD: Representació sense mandat exprés.....	293
5.2. Possibilitats de l'article 80.2 RGPD per redreçar la discriminació algorítmica.....	296
5.3. Dificultats per l'exercici de l'article 80 en relació amb la discriminació algorítmica.....	297
6. Conclusions parcials.....	300
Capítol VI. Discriminació algorítmica i PRIA.....	303
1. El principi de no-discriminació en la PRIA.....	306
1.1. Un Reglament per assegurar el bon funcionament del mercat intern de la UE.....	307

1.2. Poca consideració cap a les persones afectades pels sistemes d'IA.....	310
1.3. Habilitació del tractament de categories especials de dades	313
1.4. Una perspectiva basada en el risc	316
2. Protecció de la discriminació algorítmica segons el grup de risc en el qual estigui classificat el sistema d'IA.....	318
2.1. Sistemes prohibits.....	318
2.2. Sistemes de risc alt	321
2.3. Sistemes de risc baix o mínim	326
3. Elements de la PRIA rellevants per la protecció envers la discriminació	328
3.1. Disseny i qualitat de les dades	329
3.2. Documentació tècnica per dur a terme un plet de discriminació.....	332
3.3. Supervisió humana dels sistemes de risc alt.....	335
3.4. Base de dades pública.....	337
3.5. Espais de prova.....	338
4. No discriminació i estandardització de sistemes d'IA.....	341
4.1. Trasllat del principi de no-discriminació als estàndards creats per les organitzacions europees d'estandardització	341
4.2. Avaluació de conformitat.....	348
4.3. Marcatge "CE" de conformitat	352
4.4. Seguiment posterior a la comercialització o posada en servei del sistema d'IA.....	354
5. Conclusions Parcial.....	356
TERCERA PART. Protecció col·lectiva-institucional envers la discriminació algorítmica	361
Capítol VII. Protecció col·lectiva: l'acció de representació en matèria de discriminació algorítmica	365
1. Per què és necessària l'articulació d'un sistema efectiu de representació col·lectiva enfront conductes discriminatòries causades mitjançant sistemes algorítmics?	366
1.1. Promoció de l'acció de representació en matèria de discriminació algorítmica en la UE	367
1.2. Interessos pròpiament grupals (col·lectius o difusos), interessos (pluri)individuals homogenis i l'interès general.....	374
1.3. Legitimació activa en l'acció col·lectiva en matèria antidiscriminatòria..	379
1.4. Acció de cessació i acció de reparació.....	381
2. Normativa antidiscriminació i acció de representació en matèria de discriminació algorítmica	391
2.1. Acció de representació a la UE en matèria de discriminació	391
2.2. Acció de representació a Espanya en matèria de discriminació algorítmica	396

2.3. Avantatges de l'acció de representació en matèria de discriminació algorítmica en comparació amb l'acció individual.....	403
3. Normativa de protecció de dades i l'acció de representació en matèria de discriminació algorítmica	405
3.1. Acció de representació de protecció de dades a la UE en el context de la discriminació algorítmica	406
3.2. Acció de representació de protecció de dades a Espanya en el context de la discriminació algorítmica	408
3.3. Aportació de la normativa europea de protecció de dades a l'acció de representació per la lluita envers la discriminació algorítmica	411
4. PRIA i acció de representació en matèria de discriminació algorítmica	414
5. Conclusions parcials	416
Capítol VIII. Protecció institucional: organismes nacionals d'igualtat com a actors necessaris per fer front a la discriminació algorítmica	419
1. Els organismes d'igualtat dels estats membres de la UE com a òrgans més ben posicionats per desafiar decisions discriminatòries preses mitjançant sistemes algorítmics	420
1.1. Primera fase: àmbit d'actuació reduït i poques competències.....	421
1.2. Segona fase: ampliació de competències amb incidència directa en les facultats per fer front a la discriminació algorítmica.....	424
2. Funcions de l'organisme d'igualtat per fer front a la discriminació algorítmica	433
2.1. Funcions com a òrgan quasi-jurisdiccional	434
2.2. Funcions com a òrgan legitimat per iniciar litigis per protegir persones i col·lectius discriminats	438
2.3. Altres funcions	449
3. El cas espanyol: Autoritat independent per a la Igualtat de Tracte i la No Discriminació en el context de la discriminació algorítmica	454
3.1. L'adaptació de l'organisme d'igualtat espanyol als estàndards europeus: l'Autoritat Independent per a la Igualtat de Tracte i la No Discriminació	455
3.2. Funcions de l'Autoritat Independent amb incidència en matèria de discriminació algorítmica	457
4. Complementarietat, coordinació i cooperació entre l'Autoritat Independent, l'AESIA, l'AEPD	462
4.1. Diferències de regulació dels usos dels sistemes algorítmics en el sector públic i privat a efectes de l'actuació de les autoritats de supervisió envers la discriminació algorítmica	463
4.2. L'AESIA i la protecció envers la discriminació algorítmica	470
4.3. AEPD i la protecció envers la discriminació algorítmica.....	475
4.4. Complementarietat, coordinació i cooperació entre l'AESIA, l'AEPD i l'Autoritat Independent.....	479
5. Conclusions parcials	483

CONCLUSIONS	487
BIBLIOGRAFIA	512
BIBLIOGRAFIA WEB	554
JURISPRUDÈNCIA I ALTRES RESOLUCIONS	559
LEGISLACIÓ.....	569
ANNEX.....	579

INTRODUCCIÓ

La discriminació, tal com està definida pel Centre de terminologia “TERMCAT” és el *“tracte desfavorable, d'inferioritat o excloent que es dona a una persona o un col·lectiu per raó d'edat, sexe, raça, ètnia, cultura, ideologia, religió o qualsevol altra condició o circumstància de tipus personal o social¹, que limita o anul·la el gaudi o l'exercici dels drets fonamentals i les llibertats fonamentals en igualtat en les esferes política, econòmica, social, cultural i civil o en qualsevol altre àmbit de la vida pública”².*

Es defineix decisió únicament automatitzada en el document anomenat *“Guidelines on Automated individual decision-making and Profiling for the purposes of Regulation 2016/679”*, realitzat pel Grup de Treball de l'Article 29³, com a aquella basada només en mitjans tecnològics sense la participació de l'ésser humà⁴. Aquest tipus de decisions poden ser preses mitjançant intel·ligència artificial (també l'anomenarem al llarg del treball, IA), que s'entén com la capacitat d'un ordinador o un robot controlat per ordinador per realitzar tasques comunament associades amb éssers intel·ligents, aplicant-se freqüentment a desenvolupar sistemes dotats dels processos intel·lectuals característics dels humans, com la capacitat de raonar, descobrir significats, generalitzar o aprendre d'experiències passades⁵. Les decisions automatitzades mitjançant intel·ligència artificial seran les que abordarem específicament en aquest treball.

Les decisions automatitzades, tal com les decisions no automatitzades, causen situacions discriminatòries. L'automatització brinda eficàcia, celeritat, reducció de costos i capacitat per processar i analitzar una quantitat de dades que la ment humana no pot abastar. Tot i així, una màquina, inclús en el cas de les decisions automatitzades o l'aprenentatge

¹En aquest treball anomenarem les raons de discriminació protegides per la legislació atributs protegits, categories sensibles, característiques protegides, motius protegits o categories protegides.

²Discriminació. TERMCAT Centre de Terminologia, 2023. Disponible a: <https://www.termcat.cat/ca/diccionaris-en-linia/228/fitxa/MzU0MjAwNA==>. Darrera visita el 2 de setembre de 2023.

³Arran de l'entrada en vigor del Reglament General de Protecció de Dades, aquest grup de treball europeu independent s'anomena Comitè Europeu de Protecció de Dades (CEPD), en anglès, European Data Protection Board (EDPB). Aquest canvi es deu a que el Grup de Treball de l'Article 29 interpretava la Directiva 95/46/CE, que va quedar derogada amb l'entrada en vigor de l'RGPD.

⁴ARTICLE 29 DATA PROTECTION WORKING PARTY, *Guidelines on Automated individual decision-making and Profiling for the purposes of Regulation 2016/679*, 17/EN, WP 251, Brussels, 6 February 2018, pàg. 8.

⁵Artificial Intelligence. Britannica, 2023. Disponible a: <https://www.britannica.com/technology/artificial-intelligence>. Darrera visita el 2 de setembre de 2023.

automàtic, segueix unes directrius donades, per tant, l'elecció sobre quines dades s'utilitzen, com es tracten i les conclusions a les que s'ha d'arribar o els objectius que té l'acció, estan, en primera instància, estipulats per persones. Això comporta que aquests sistemes o bé puguin integrar o mantenir estigmatitzacions i prejudicis que afectin negativament membres de certs grups històricament marginats únicament perquè han après a fer-ho⁶, o bé que les persones que participen en el seu disseny o garants del seu funcionament puguin introduir els seus propis biaixos o traslladin erròniament legislacions o polítiques a seguir, causant que el sistema demani coses diferents a aquelles realment establertes⁷.

Les decisions automatitzades mitjançant sistemes algorítmics⁸ tenen la capacitat, a causa de la seva creixent incidència en la vida dels éssers humans, d'ocasionar riscos pels drets i llibertats fonamentals de les persones. Dins el marc regulador europeu, la protecció que es brinda als individus en relació a aquest tipus de decisions es troba de manera principal al Reglamento (UE) 2016/679 del Parlamento Europeo y del Consejo de 27 de abril de 2016 relativo a la protección de las personas físicas en lo que respecta al tratamiento de datos personales y a la libre circulación de estos datos⁹ (d'ara en endavant, RGPD), en l'article 22, que estipula el dret a no ser objecte d'una decisió basada únicament en el tractament automatitzat, que produeixi efectes jurídics o tingui una afectació significativa de manera similar¹⁰.

En el pla semàntic, com ens centrarem en el debat jurídic i no en la precisió tècnica, utilitzarem, per referir-nos al mateix concepte, les expressions “sistemes d'intel·ligència artificial”, “sistemes algorítmics”, “sistemes de presa de decisions mitjançant algoritmes” i “decisiones automatitzades mitjançant sistemes algorítmics”. També utilitzem frases com “efectes de la intel·ligència artificial” tal com si la intel·ligència artificial fos una entitat que actués per si mateixa. Així i tot, els sistemes d'intel·ligència artificial no es creen

⁶KAMINSKI, M.E., “Binary Governance: Lessons from the GDPR's Approach to Algorithmic Accountability”, *Southern California Law Review*, vol. 92, 2019, pàg. 1542.

⁷LUNDBERG, E., *Automated decision-making vs indirect discrimination. Solution or aggravation?*, Umeå University, Umeå, 2019, pàg. 26.

⁸ Al llarg del treball quan ens referim a decisions automatitzades mitjançant sistemes algorítmics o sistemes d'IA en referim als casos en que aquestes s'apliquen en contextos socials i poden produir discriminació.

⁹ Reglamento (UE) 2016/679 del Parlamento Europeo y del Consejo de 27 de abril de 2016 relativo a la protección de las personas físicas en lo que respecta al tratamiento de datos personales y a la libre circulación de estos datos y por el que se deroga la Directiva 95/46/CE (Reglamento general de protección de datos). DOUE, L 119/1, 4.5.2016.

¹⁰Article 22 de l'RGPD.

espontàniament i no tenen per ells mateixos efectes adversos en els drets humans, sinó que els tenen en tant s'implementen i s'apliquen a la interacció humana.

La problemàtica que se'ns presenta resideix en els riscos afegits per la discriminació que trobem en les decisions automatitzades, i les mancances que s'observen en el marc regulador europeu actual, que impossibiliten protegir d'una manera efectiva les persones enfront les discriminacions que poden ocasionar les decisions automatitzades. Específicament, centrarà la nostra atenció la dificultat que creen aquest tipus de decisions per tal que la persona suposadament discriminada aporti un cas que es pugui comparar amb el seu (és a dir, com poder provar que ha existit una discriminació amb absència pràctica de *tertium comparationis*), i si a través d'un sistema de protecció col·lectiva-institucional es pot garantir millor el dret de les persones que es veuen discriminades.

En conseqüència, en aquest treball ens plantegem cinc qüestions de recerca:

1. Les decisions automatitzades augmenten les possibilitats de discriminació?
Per què?
2. El marc regulador europeu antidiscriminatori i el marc regulador europeu de protecció de dades són suficients per fer front a les discriminacions arran de les decisions automatitzades?
3. Algunes garanties de la regulació antidiscriminació i de protecció de dades podrien ser vàlides per protegir les persones de la discriminació algorítmica donades certes condicions?
4. Ofereix noves garanties el nou marc europeu per regular la intel·ligència artificial?
5. És necessària la promoció de la representació col·lectiva per fer front a la discriminació algorítmica? Cal una institució pública que coordini la lluita envers la discriminació algorítmica?

Per poder respondre a aquestes qüestions, la metodologia emprada és la de l'anàlisi de la doctrina especialitzada, la regulació i la jurisprudència, per determinar l'estat de la qüestió i la descripció dels problemes que plantegen les decisions automatitzades mitjançant sistemes algorítmics a la protecció contra les discriminacions.

Per fer-ho, el treball es divideix en tres parts. La primera part se centra en els límits que existeixen en el marc regulador europeu actual per fer front a la discriminació ocasionada mitjançant decisions automatitzades a través de sistemes algorítmics (que també anomenarem discriminació algorítmica). Cal estudiar, per consegüent, com poden causar discriminació les decisions automatitzades, i la regulació que s'ha adoptat en l'àmbit de la Unió Europea (que també anomenem al llarg del treball, UE) per protegir les persones d'aquesta discriminació. Es fa palès que les decisions automatitzades preses mitjançant sistemes algorítmics poden tenir efectes discriminatoris, i que la regulació europea analitzada presenta punts febles per poder lluitar envers la discriminació produïda.

La segona part se centra, en contrapartida, en quins elements de les regulacions europees estudiades en la primera part del treball poden, donades certes condicions, ser d'utilitat per oferir certa protecció, combinades amb les garanties establertes en la Propuesta de Reglamento del Parlamento europeo y del Consejo, por el que se establecen normas armonizadas en materia de inteligencia artificial (ley de inteligencia artificial) y se modifican determinados actos legislativos de la Unión¹¹ (d'ara en endavant, PRIA). Queda patent que la regulació antidiscriminació europea i l'RGPD, sota certes condicions, i en combinació amb les garanties establertes en la PRIA, poden oferir certes eines a les persones per defensar-se envers la discriminació algorítmica. Tot i així, aquestes garanties no són suficients.

Finalment, en la tercera part, plantegem un sistema de protecció col·lectiva-institucional, que reuneix els punts forts de les tres normatives analitzades al llarg del treball, per tal que els estats membres de la UE protegeixin els col·lectius perjudicats per discriminació algorítmica, i alhora proporcionin les eines suficients a les persones que s'han vist discriminades per poder desafiar decisions automatitzades mitjançant sistemes algorítmics presumptivament discriminatòries.

¹¹ COMISIÓN EUROPEA, *Propuesta de Reglamento del Parlamento europeo y del Consejo, por el que se establecen normas armonizadas en materia de inteligencia artificial (ley de inteligencia artificial) y se modifican determinados actos legislativos de la Unión*, COM(2021) 206, Bruselas, 24 de abril de 2021.

PRIMERA PART. Límits de la regulació europea per protegir les persones de la discriminació causada per decisions automatitzades

La confiança en la intel·ligència artificial i els algoritmes deriva de la creença que els agents no humans no tenen prejudicis, i que les seves decisions no estan afectades per ideologies. De fet, els algoritmes tenen tants prejudicis com les persones que els han entrenat, però d'una manera més opaca i que fa més difícil la rendició de comptes¹². A la vegada, a través dels seus efectes a gran escala, també poden oferir oportunitats per visualitzar, mesurar, detectar i corregir millor els biaixos discriminatoris si s'estableix la legislació adequada¹³.

Els objectius que ens proposem en aquesta primera part del treball són, en primer lloc, determinar com poden discriminar les decisions preses mitjançant sistemes algorítmics i, en segon lloc, estudiar quins elements limiten el marc regulador europeu potencialment aplicable.

Interessa exposar certes precisions en relació al contingut abans d'abordar el primer capítol. Les decisions automatitzades que empren intel·ligència artificial es poden dur a terme mitjançant algoritmes no predictius o mitjançant algoritmes predictius. Seguirem l'exemple de Huergo Lora a "*La regulación de los algoritmos*", basat en el sistema d'acreditació de professorat universitari. Qui realitza la norma decideix, fent un judici de valor, quines són les característiques d'un/a professor/a ideal i puntua les persones aspirants en el grau en què les posseeixen. Un algoritme no predictiu calcularà les puntuacions en base a les dades curriculars (publicacions, docència, etc.). En canvi, un algoritme predictiu, analitzarà les dades del professorat considerat millor o més productiu d'acord amb els criteris prèviament establerts; seguidament, determinarà quines característiques formatives o curriculars estan associades a aquest nivell d'excel·lència; i

¹²NOTO LA DIEGA, G., "Against the Dehumanisation of Decision-Making Algorithmic Decisions at the Crossroads of Intellectual Property, Data Protection, and Freedom of Information", *JIPITEC- Journal of Intellectual Property, Information Technology and E-Commerce Law*, vol. 9, núm. 3, 2018, pàg. 33.

¹³ GERARDS, J. i XENIDIS, R., *Algorithmic Discrimination in Europe: Challenges and Opportunities for Gender Equality and Non-Discrimination Law*, European Commission, Brussels, 2021, pàg. 10.

finalment, identificarà d'entre les persones aspirants, aquelles que reuneixen aquestes característiques i, per tant, es pot preveure que arribaran al nivell establert¹⁴.

Els algoritmes predictius són la base de l'aprenentatge automàtic. Però, que és l'aprenentatge automàtic (en tant que eina que pot ser utilitzada en la presa de decisions automatitzades)? Alpaydin observà que els sistemes que programaven els investigadors per resoldre qüestions tenien dos desavantatges, en primer lloc, que les normes lògiques en els sistemes no sempre encaixaven la realitat canviant del món (en la vida quotidiana les coses no són certes o falses pròpiament); i en segon lloc, les persones investigadores havien de proveir el coneixement (les respostes) per introduir-lo en el sistema, un procés que costava molts diners i temps¹⁵. Arran de l'aprenentatge automàtic, el coneixement va passar de proveir-se per part de persones expertes a ser extret pel propi sistema en base a exemples i experiència en forma de dades (concretant-li una tasca i alimentant-lo amb una gran quantitat de dades per usar-les com a exemples de com aquesta tasca es pot aconseguir o de quines dades extreure'n patrons)¹⁶. Els recents avenços de la intel·ligència artificial es deuen, sobretot, al desenvolupament d'aquesta tècnica. Una subclasse de l'aprenentatge automàtic és l'aprenentatge profund (Deep learning). Dins d'aquest es pot distingir entre l'aprenentatge supervisat, és a dir, quan les dades d'entrada que utilitza el sistema són etiquetades per persones, i l'aprenentatge no supervisat, és a dir, quan les dades d'entrada no són etiquetades per persones i és un algoritme que realitza la seva pròpia classificació¹⁷.

També, és rellevant diferenciar entre els conceptes de discriminació directa i discriminació indirecta. A grans trets i malgrat s'abordarà amb més precisió en els apartats pertinents, la discriminació directa és aquella situació en la que una persona és tractada de manera menys favorable del que ha sigut o vagi a ser tractada una altra persona en una situació comparable a causa d'una categoria protegida (per exemple l'ètnia o el gènere)¹⁸ i, en canvi, la discriminació indirecta es podria definir com aquella situació en la qual una

¹⁴ HUERGO LORA, A.J., "Una aproximación a los algoritmos desde el derecho administrativo", a HUERGO LORA, A.J., (dir.) i DÍAZ GONZÁLEZ, G. M. (coord.), *La regulación de los algoritmos*, Pamplona, Aranzadi, 2020, pàg. 38.

¹⁵ALPAYDIN, E., *Machine learning: the new AI*, Massachusetts, The MIT Press, 2016, pàg. 51.

¹⁶THE ROYAL SOCIETY, *Machine learning: the power and promise of computers that learn by exemple*, London, April 2017, pàg.19.

¹⁷OBSERVATORI D'ÈTICA EN INTEL·LIGÈNCIA ARTIFICIAL DE CATALUNYA (OEIAC), *Intel·ligència artificial, ètica i societat: Una mirada i discussió a través de la literatura especialitzada i d'opinions expertes*, Girona, 2021, pàg. 16 i 17.

¹⁸REY MARTÍNEZ, F., *Derecho Antidiscriminatorio*, Navarra, Aranzadi, 2019, pàg. 56.

pràctica que aparentment és neutra acaba situant en una posició de desavantatge particular a persones que es poden identificar amb alguna característica protegida en comparació amb persones que no ostenten la categoria protegida¹⁹.

Després d'aquestes apreciacions, a través dels tres capítols que seguidament plantejem, pretenem donar resposta a les dues primeres qüestions de recerca d'aquest treball, referents a si les decisions automatitzades augmenten les possibilitats de discriminació i per què, i si el marc regulador europeu antidiscriminació i el marc regulador europeu de protecció de dades són suficients per protegir les persones de manera efectiva.

En el capítol I, estudiem quins són els elements que fan que les decisions automatitzades mitjançant sistemes algorítmics siguin discriminatòries; en el capítol II analitzem els límits de la normativa antidiscriminació europea per abordar les discriminacions arran de decisions automatitzades mitjançant sistemes algorítmics, i finalment, en el capítol III ens centrem en les limitacions respecte la protecció de les persones envers les discriminacions que trobem en l'element principal regulador dins el marc legislatiu europeu actual de les decisions automatitzades, l'RGPD.

¹⁹TJUE (Sala Tercera), sentència de 24 de febrer de 2022, C-389/20, CJ contra Tesorería General de la Seguridad Social (TGSS), par. 40.

Capítol I. Per què les decisions automatitzades poden ser discriminatòries?

En el present capítol analitzem com les decisions automatitzades mitjançant sistemes algorítmics poden ocasionar entrebancs afegits en la lluita per la igualtat i la no discriminació. Principalment abordem la primera qüestió que ens plantejem en aquest treball, concretament, si les decisions automatitzades augmenten les possibilitats de discriminació, i si és així, per què passa.

L'esforç requerit per identificar les causes de resultats injustos i discriminatoris poden esperar-se exponencialment més grans, com exponencialment més complexos seran els sistemes emprats per assistir en els processos de presa de decisions. Aquest escenari subratlla la necessitat que aquests sistemes siguin transparents i estiguin subjectes a rendició de comptes. Adoptar sistemes opacs de decisió quan els resultats poden afectar significativament tant la vida de les persones, com les seves opinions i percepcions, pot derivar en formes d'estigma algorítmic, un escenari molt perillós tenint en compte que aquelles persones estigmatitzades potser mai en seran conscients i aquella qui estigmatitzi ho farà (conscientment o inconscientment) a través d'una màquina²⁰. En conseqüència, aquesta nova eina, l'ús de la qual s'està incrementant vertiginosament, planteja nous reptes als ja existents, de manera que cal abordar per què i de quines maneres les decisions automatitzades poden ser discriminatòries.

Per fer-ho, aquest capítol consta de tres parts ben diferenciades, estructurades segons un anàlisi deductiu, d'allò més general a allò més particular. A mode introductori, comencem relacionant els conceptes d'igualtat, equitat i discriminació amb els sistemes algorítmics (1), explicant posteriorment com ocorre la discriminació algorítmica (2) i finalment, localitzant quins riscos i problemes afegits podem trobar en el seu funcionament per lluitar contra la igualtat i la no discriminació, a saber, els biaixos algorítmics, l'opacitat inherent als sistemes d'intel·ligència artificial, la personalització i la magnificació dels efectes discriminatoris (3). Acabem el capítol amb un apartat de conclusions parcials (4).

²⁰LEPRI, B., STAIANO, J., SANGOKOYA, D., LETOUZÉ, E. i OLIVER, N., "The tyranny of data? the bright and dark sides of data-driven decision-making for social good", a CERQUITELLI T., QUERCIA D., i PASQUALE F., (eds.), *Transparent Data Mining for Big and Small Data*, Switzerland, Springer, 2017, pàg. 12.

1. Conceptes introductoris

En aquest apartat, com ja hem anunciat, relacionem els conceptes d'igualtat (1.1), equitat (1.2) i discriminació (1.3), amb els sistemes algorítmics.

1.1. La igualtat i els sistemes algorítmics

Amb l'objectiu d'entendre millor la rellevància de la concepció de la igualtat per abordar la discriminació, en aquest apartat definirem igualtat, introduïrem el concepte des de la vessant jurídica, presentarem la diferència entre igualtat formal i igualtat substantiva i explicarem la perspectiva de la UE i la seva posició enfront el desenvolupament de la intel·ligència artificial.

La igualtat es pot definir com la conformitat d'alguna cosa amb una altra en naturalesa, forma, qualitat o quantitat²¹. Com a principi general del dret, exigeix que s'atorgui davant situacions iguals el mateix tracte i davant situacions desiguals s'afavoreixi un tracte diferent²². Tot i així, anant més enllà de la mera proclamació, la igualtat en dret és raonabilitat de la diferència jurídica de tracte²³, de manera que no es pot entendre com una obligació que tots els individus siguin tractats de la mateixa manera, ni tampoc, contràriament, que es permeti tota diferenciació de tracte. El poder legislatiu, en regular qualsevol assumpte, estableix, normalment, diverses diferències de tracte entre els destinataris, en funció de les circumstàncies. De fet, allò que és normal en dret, com en la naturalesa, és la desigualtat de tracte, no la igualtat. La igualtat és llavors, una tècnica de

²¹Igualdad. Diccionario de la Real Academia Española en línea. 2023. Disponible a: <https://dle.rae.es/igualdad>. Darrera visita el 2 de setembre de 2023.

²²Principio de igualdad. Diccionario jurídico de derecho en línea. Enciclopedia Jurídica. 2023. Disponible a: <http://www.encyclopedia-juridica.com/d/principio-de-igualdad/principio-de-igualdad.htm>. Darrera visita el 2 de setembre de 2023.

²³MARTÍNEZ GARCÍA, J. I., "El principio de igualdad y la producción de diferencias en el derecho" , a MINISTERIO DE JUSTÍCIA i CENTRO DE PUBLICACIONES, (eds.), *El principio de igualdad en la Constitución española: XI Jornadas de Estudio*, vol. 1, 1991, pàg. 539 i s. Vegeu: Tribunal Constitucional (Pleno), sentència núm. 223/1981, de 2 de julio de 1981, FJ 3.

control, un criteri que mesura el grau de desigualtat jurídicament administrable, així que, estem en realitat, parlant de quanta desigualtat és tolerable²⁴.

A Espanya, el Tribunal Constitucional (d'ara en endavant, TC), ha especificat que l'article 14 de la Constitució (principi d'igualtat davant la llei i prohibició de discriminació) es desplega en diferents nivells, concretament, la igualtat davant la llei i la igualtat en la llei. La sentència 103/1983, de 22 de novembre, del TC, disposa que l'article 14 de la Constitució imposa un límit a la potestat del poder legislatiu i atorga un dret subjectiu, en fer que les normes legals no creïn entre la ciutadania situacions desiguals o discriminatòries i obrint la possibilitat de posar en marxa els mecanismes jurídics idonis per restablir la igualtat trencada, respectivament²⁵. En la sentència 75/1983, de 3 d'agost, també del TC, s'exposa que la configuració com un dret subjectiu dels/de les ciutadans/es, evita els privilegis i les desigualtats discriminatòries entre ells, sempre que aquestes es trobin dins de les pròpies situacions de fet a les que ha de correspondre un tracte jurídic igual²⁶. Per tant, una diferenciació no és discriminatòria ni contrària al principi d'igualtat si el seu contingut es recolza en una situació diferenciada que rep un tractament singular basat en fonaments raonables segons criteris de valor acceptats amb generalitat²⁷.

Sovint s'oposa la igualtat "formal" o "jurídica" a la igualtat "real" o "d'oportunitats"²⁸. La igualtat de tracte jurídic demostra moltes deficiències per abordar molts conflictes presents en la interacció entre éssers humans. A ella s'oposa la igualtat material o substantiva, és a dir, l'establiment de diferències de tracte per tal d'aconseguir una igualtat més real i efectiva.

²⁴REY MARTÍNEZ, F., *Derecho Antidiscriminatorio*, op. cit., 2019, pàg. 26; SORIANO ARNANZ, A., *Data Protection for the prevention of algorithmic discrimination. Protecting from discrimination and other harms caused by algorithms through privacy in the EU: possibilities, shortcomings and proposals*, Pamplona, Aranzadi, 2021, pàg. 242. L'autora, en referència a l'afectació de les decisions preses mitjançant sistemes algorítmics sobre el dret a la igualtat i a la no discriminació determina que serà necessari establir quin grau d'igualtat es considerat acceptable, la qual cosa pot suposar establir certes restriccions per l'ús privat d'algoritmes que no serien considerats en un context no automatitzat.

²⁵ Tribunal Constitucional (Pleno), sentència 103/1983, de 22 de noviembre de 1983, FJ 3.

²⁶ Tribunal Constitucional (Pleno), sentència 75/1983, de 3 de agosto de 1983, FJ 2.

²⁷ *Ídem*, FJ 7.

²⁸Per exemple, en la Constitució espanyola de 1978, es recull en l'article 14 la igualtat formal o jurídica i en l'article 9.2 la igualtat "real i efectiva" d'individus i grups. Transcrivim aquí l'article 9.2 de la Constitució Espanyola de 1978: "*Corresponde a los poderes públicos promover las condiciones para que la libertad y la igualdad del individuo y de los grupos en que se integra sean reales y efectivas; remover los obstáculos que impidan o dificulten su plenitud y facilitar la participación de todos los ciudadanos en la vida política, económica, cultural y social*"; i l'article 14 de la Constitució Espanyola de 1978: "*Los españoles son iguales ante la ley, sin que pueda prevalecer discriminación alguna por razón de nacimiento, raza, sexo, religión, opinión o cualquier otra condición o circunstancia personal o social*".

La dificultat per al reconeixement i protecció dels drets a la igualtat i a la no discriminació ve causada principalment per les tensions que existeixen amb el lliure desenvolupament de l'autonomia de la voluntat i l'eficiència entesa des d'una perspectiva econòmica²⁹. Aquests dos conceptes estan estretament lligats a la llibertat, però s'ha de tenir en compte que aquesta només existeix quan les persones tenen *de facto* possibilitats d'escollir, és a dir, tenen opcions. Quan es neguen oportunitats pel fet de pertànyer a un grup socialment marginat, s'està negant l'oportunitat d'escollir. Per tant, una mesura que aparentment limita la llibertat, pot servir per garantir la llibertat d'aquelles persones que a causa de tenir certes característiques inherents a la seva persona infravalorades o mal valorades dins la societat parteixen en desavantatge en molts àmbits de la vida social, econòmica i jurídica³⁰. Aquí parlem de la igualtat substantiva i la igualtat d'oportunitats, entesa aquesta última en l'àmbit sociològic com la circumstància social en la qual la probabilitat d'èxit personal no està influïda per l'origen social³¹ (tractarem aquesta qüestió amb més profunditat en l'apartat 2 del capítol II).

L'ordenament jurídic europeu en relació amb la igualtat i la no discriminació ha adoptat una perspectiva bàsicament formal, reconeixent una especial protecció enfront la discriminació en base a certes categories protegides³², considerant que aquestes són, en principi, immutables o pertanyen a l'espai inherent a la dignitat humana³³. Sota els paràmetres de prohibició en base a categories protegides també s'estan perfilant les regulacions d'eines tecnològiques de les quals es vol regular la seva implementació, com per exemple els mecanismes de reconeixement facial, tenint en compte que en les guies proposades pel Comitè Consultiu per la protecció de les persones en relació al tractament

²⁹OKUN, A. M., *Equality and Efficiency: The Big Tradeoff*, Washington, D.C., Brookings Institution Press, 1975, pàg. 12.

³⁰KEREN, H., "We insist! Freedom now: Does contract doctrine have anything constitutional to say?", *Michigan Journal of Race Law*, vol. 11, núm. 1, 2005, pàg. 174-175.

³¹Igualtat d'oportunitats. Diccionari en línia. Enciclopèdia Catalana. 2023. Disponible a: <https://www.enciclopedia.cat/ec-gec-0517880.xml>. Darrera visita el 2 de setembre de 2023.

³²Tant l'article 21 de la Carta de Drets Fonamentals de la UE com l'article 14 del Conveni Europeu de Drets Humans utilitzen la mateixa fórmula: enumeren certes categories protegides i deixen la llista oberta per incorporar-ne d'altres. En canvi, el dret derivat de la UE desenvolupa la seva protecció en base a certes categories protegides, a saber, el sexe, l'origen racial o ètnic, l'edat, la discapacitat, la religió o creences i l'orientació sexual. En tot cas, la protecció s'articula en base a categories protegides.

³³SORIANO ARNANZ, A., "Decisiones automatizadas y discriminación: aproximación y propuestas generales", *Revista General de Derecho Administrativo*, núm. 56, 2021, pàg. 9.

de dades personals de la Convenció 108, es planteja la prohibició del seu ús quan la seva única finalitat sigui determinar el color de pell d'una persona, el gènere, l'edat, etc.³⁴

Aquesta perspectiva està construïda en base a unes estructures el fonament de les quals ha estat mantenir la llibertat econòmica establerta històricament des de la perspectiva d'aquelles persones que detenien el poder, en conseqüència, la protecció de la igualtat, tot i reconeguda, manca de la profunditat requerida per tallar d'arrel la discriminació sistèmica existent a la societat, protegint-se, per defecte, en el marc de la UE, la llibertat d'empresa i situant-se en un pla secundari la protecció dels drets fonamentals a la igualtat i a la no discriminació³⁵. Partint d'aquesta base, des del nostre punt de vista, una protecció de la igualtat i la no discriminació feble ha de fer front a un gran producte que pot arribar a produir grans beneficis, com és la intel·ligència artificial, de manera que parteix en desavantatge. Per exemple, les decisions automatitzades mitjançant sistemes algorítmics es van presentar inicialment com una eina més objectiva i per tant, més respectuosa amb la igualtat que la presa de decisions duta a terme per éssers humans, ja que s'entenia que era una màquina la que decidia, inherentment objectiva, i no una persona, inherentment subjectiva. Aquesta concepció es va posar en entredit en evidenciar-se que les màquines podien contenir els mateixos biaixos i prejudicis que les persones que les havien ideat, afegint-se que l'impacte era molt més gran en poder prendre decisions analitzant dades massivament.

Tractem en el següent apartat l'equitat, donada la rellevància que està adquirint la perspectiva de la igualtat real i de fet com a pilar per abordar la regulació dels sistemes algorítmics.

1.2. L'equitat i els sistemes algorítmics

En aquest apartat definim l'equitat i abordem, per una banda, les raons per les quals hi ha una divisió doctrinal entre si l'equitat és una bona manera d'abordar la discriminació

³⁴CONSULTATIVE COMMITTEE OF THE CONVENTION FOR THE PROTECTION OF INDIVIDUALS WITH REGARD TO AUTOMATIC PROCESSING OF PERSONAL DATA, *Guidelines on Facial Recognition*, T-PD(2020)03rev4, Strasbourg, 28 January 2021, pàg. 5.

³⁵SORIANO ARNANZ, A., "Decisiones automatizadas y discriminación...", *op. cit.*, 2021, pàg. 10.

algorítmica i donar solucions alternatives a les plantejades fins ara o no (1.2.1). Plantegem, seguidament, exemples pràctics d'equitat en sistemes algorítmics i les seves mancances (1.2.2).

1.2.1. L'equitat (fairness) com a línia teòrica que emmarca la discriminació algorítmica

L'equitat és l'adaptació de la norma jurídica a les particularitats de cada cas concret, per evitar una interpretació rígida dels preceptes legals d'acord amb l'esquema general de les normes jurídiques³⁶. S'ha anat contemplant en diversos països la idea d'equitat com a superació de la simple "igualtat formal" i equivalent, a la "igualtat real o d'oportunitats". Significa anivellar les condicions de vida d'aquelles persones socialment vulnerables, un judici de valor que pren partit no per tractar a totes les persones igual, sinó per tractar millor a aquelles que estan pitjor; fugint de la simetria de la idea formal d'igualtat. La noció d'equitat és més aviat política, no hi ha una definició jurídica. És igualtat substantiva i per això, objectiu central del dret antidiscriminatori³⁷. Si el dret no tracta de manera diferent i millor a aquelles persones que estan de fet pitjor, consolida i beneeix la subordinació³⁸.

La Resolución del Parlamento Europeo, de 14 de marzo de 2017, sobre las implicaciones de los macrodatos en los derechos fundamentales: privacidad, protección de datos, no discriminación, seguridad y aplicación de la ley³⁹, en el seu apartat 22, que versa sobre l'impacte discriminatori dels sistemes d'algoritmes i conjunts de dades, es refereix a l'equitat com a criteri que ha de guiar l'examen de les prediccions basades en anàlisi de dades. La comunitat investigadora de diferents disciplines ha incrementat l'atenció en promoure, entendre i abordar la lluita per la igualtat i la no discriminació en la intel·ligència artificial a través de l'equitat⁴⁰, és a dir, mitjançant la intervenció en els

³⁶Equitat. Diccionari en línia. Enciclopèdia Catalana. 2023. Disponible a: <https://www.enciclopedia.cat/ec-gdlc-e00054409.xml>. Darrera visita el 2 de setembre de 2023.

³⁷REY MARTÍNEZ, F., *Derecho Antidiscriminatorio*, op. cit., 2019, pàg. 31.

³⁸*Ídem*, 46.

³⁹PARLAMENTO EUROPEO, *Resolución del Parlamento Europeo, de 14 de marzo de 2017, sobre las implicaciones de los macrodatos en los derechos fundamentales: privacidad, protección de datos, no discriminación, seguridad y aplicación de la ley*, 2016/2225(INI), Bruselas, 2017.

⁴⁰GERARDS, J. i XENIDIS, R., *Algorithmic Discrimination in Europe...*, op. cit., 2021, pàg. 48. El principi d'igualtat en el context de la UE té un àmbit d'aplicació més restrictiu que el concepte d'equitat entès pels

resultats o en el procés de presa de decisió mitjançant sistemes algorítmics per fer que aquests siguin més justos i no produeixin desavantatges per grups específics vulnerables, dipositant molts esforços en desenvolupar marcs d'equitat o models algorítmics per alleujar els biaixos⁴¹. També, des de 2018 es duu a terme anualment l'ACM Conference on Fairness, Accountability and Transparency, una sèrie de conferències acadèmiques sobre ètica i sistemes informàtics en la qual un dels temes principals a tractar és l'equitat.

No obstant això, no hi ha una opinió unànime sobre la seva adequació com a perspectiva per abordar la discriminació en els sistemes algorítmics, considerant l'existència de severes limitacions del concepte com a mètode per avaluar l'impacte social dels algoritmes. En primer lloc, estudis existents sobre l'equitat en els sistemes algorítmics són bàsicament normatius (de caràcter prescriptiu), per tant, únicament defineixen com s'haurien de dur a terme decisions justes assumint que existeix consens social sobre què constitueix una decisió justa⁴². Podem subratllar, per exemple, com els algoritmes considerats justos poden reforçar la discriminació, ja que s'obvia l'opressió estructural, és a dir, l'algoritme "just" fa esdevenir els productes d'una discriminació històrica en fets neutres i empírics, reforçant aquesta discriminació i castigant aquells que han estat subjectes a aquestes circumstàncies en primer lloc⁴³. S'ha destacat llavors que l'equitat algorítmica emmiralla algunes de les tendències més problemàtiques del discurs antidiscriminació, a saber, el fracàs d'adreçar la lògica jeràrquica que produeix persones avantatjades i desavantatjades inicialment⁴⁴.

científics de la computació. El principi d'igualtat únicament imposa obligacions en relació a la protecció dels sis atributs següents: sexe, raça o origen ètnic, discapacitat, orientació sexual, religió o creença i edat (art.19 TFUE). Prohibeix de manera general el tracte o impacte de manera desfavorable (l'acció positiva es permet però no existeix una obligació legal als estats membres a adoptar-la), en canvi, els procediments d'equitat algorítmica poden ser més amplis i anar més lluny que aquesta concepció.

⁴¹Per exemple: CHOULDECHOVA, A., "Fair prediction with disparate impact: A study of bias in recidivism prediction instruments", *Big data*, vol. 5, núm. 2, 2017, pàg. 153–163; LAHOTI, P., GUMMADI, K. P. i WEIKUM, G., "Learning individually fair data representations for algorithmic decision making", *IEEE, 35th International Conference on Data Engineering (ICDE)*, 2019, pàg. 1334–1345; LAHOTI, P., GUMMADI, K. P. i WEIKUM, G., "Operationalizing individual fairness with pairwise fair representations", *Proceedings of the VLDB Endowment*, vol.13, núm. 4, 2019, pàg. 506–518.

⁴²GRGIĆ-HLAČA, N., REDMILES, E. M., GUMMADI, K. P. i WELLER, A., "Human perceptions of fairness in algorithmic decision making: A case study of criminal risk prediction", *Proceedings of the 2018 World Wide Web Conference on World Wide Web*, 2018, pàg. 903.

⁴³GREEN, B., "The False Promise of Risk Assessments: Epistemic Reform and the Limits of Fairness", *FAT* '20: Proceedings of the 2020 Conference on Fairness, Accountability, and Transparency*, 2020, pàg. 595-601.

⁴⁴HOFFMANN, A. L., "Where fairness fails: data, algorithms and the limits of antidiscrimination discourse", *Information, Communication and Society*, vol. 22, núm. 7, 2019, pàg. 901.

En segon lloc, també esdevé conflictual el fet que una perspectiva contextual apropiada de l'equitat que capturi l'essència dels conceptes filosòfics rellevants pot dependre de factors que no es troben en les dades. Aquestes dades que falten poden incloure les característiques protegides de les persones afectades però també informació rellevant per valorar-les -com les circumstàncies socioeconòmiques, l'experiència i el desenvolupament personal i les relacions entre elles. Intentos de treure conclusions a partir de les dades de formació de sistemes i de protegir les persones a partir de categories de dades semblen mètodes poc proclius a proporcionar justícia de la manera que les qüestions de justícia sorgeixen en les vides idiosincràtiques i en contextos socials diferents⁴⁵. No es pot obviar el conflicte inherent entre els procediments d'igualtat formal i els resultats d'igualtat substantiva en una societat desigual⁴⁶.

Finalment, en tercer lloc, s'ha fet palès que les mesures d'equitat individual poden entrar en contradicció amb les mesures d'equitat grupal. Per exemple, l'equitat individual (assegurar que les persones amb perfils similars seran tractades de la mateixa manera tenint en compte les característiques de la persona per adoptar una decisió), i l'equitat grupal (assegurar que el procés de decisió no afavoreix a un grup determinat arbitràriament tenint en compte la pertinença d'un individu a un grup per adoptar una decisió) poden ser intrínsecament incompatibles⁴⁷.

1.2.2. Mancances de l'aplicació pràctica de l'equitat en els sistemes algorítmics

Com a perspectiva per abordar la discriminació algorítmica, exemplificarem en aquest apartat com es pot dur a terme l'equitat a la pràctica i quines mancances s'observen en les mesures concretes aplicades per fer els processos de presa de decisió mitjançant algoritmes més equitatius.

Defensen Xenidis i Senden que la recerca de la igualtat substantiva i l'establiment d'una regulació més òptima per tractar la discriminació algorítmica s'ha d'enfocar explorant i

⁴⁵BINNS, R., "Fairness in machine learning: Lessons from political philosophy", *Proceedings of the 1st Conference on Fairness, Accountability and Transparency*, PMLR, vol. 81, 2018, pàg. 157.

⁴⁶GREEN, B., "The False Promise...", *op. cit.*, 2020, pàg. 602.

⁴⁷ INSTITUT MONTAIGNE, *Algorithmes: contrôle des biais S.V.P.*, Paris, Mars 2020, pàg. 36.

desenvolupant “la igualtat des del disseny” (tractarem aquesta qüestió amb més profunditat en l’apartat 4.3 del capítol V), una perspectiva de l’enginyeria algorítmica que caldria adoptar en cada pas del disseny algorítmic i procés de construcció, desenvolupament i posada en funcionament del sistema⁴⁸. Es preveuen situacions de classificació en les quals és desitjable ajustar o maquillar una mètrica, com per exemple en les admissions universitàries a Estats Units d’Amèrica del Nord⁴⁹, afegint cert nombre de punts a l’alumnat que formi part d’algun grup social vulnerable⁵⁰.

També s’ha plantejat l’exclusió de les categories personals sensibles de les dades emprades per la presa de decisions, malgrat s’ha reconegut com a una perspectiva insuficient per fer el procés més just ja que aquestes dades poden ser implícites en altres atributs no exclosos⁵¹, reflexió en la mateixa línia que seguirem en altres apartats d’aquest treball relatiu al fet que quan ens referim a l’anàlisi massiva de dades, totes les dades poden esdevenir dades personals.

També es proposa modificar o seleccionar les dades de manera que el resultat satisfaci l’exigència d’equitat seleccionant dades de mostra extres dels grups minoritaris o duplicant les que ja estan disponibles per assegurar-nos que un classificador no representa excessivament l’estatus de minoria d’algun grup. Però cal remarcar que és perillós aquest procediment ja que aquestes actuacions poden esbiaixar el procés d’aprenentatge del sistema o acabar adequant el model a dades no representatives⁵².

Finalment, es poden dur a terme accions positives com a mesures per la realització efectiva d’un marc jurídic d’igualtat substantiva, mesures a les quals es fa referència per exemple en les directives antidiscriminació europees⁵³ o al Tractat de Funcionament de

⁴⁸ XENIDIS, R. i SENDEN, L., “EU Non-discrimination law in the era of artificial intelligence: mapping the challenges of algorithmic discrimination”, a BERNITZ, U., GROUSSOUT, X., PAJU, J. i DE VRIES, S., (eds.), *General Principles of EU Law and the EU Digital Order*, Alphen aan den Rijn, Kluwer Law International, 2020, pàg. 179.

⁴⁹ El sistema d’accés a la universitat estatunidenc no es basa en un únic examen nacional d’accés, sinó que cada universitat estableix els seus propis estàndards o normatives d’admissió.

⁵⁰ BINNS, R., “On the Apparent Conflict Between Individual and Group Fairness”, *FAT* '20: Proceedings of the 2020 Conference on Fairness, Accountability, and Transparency*, 2020, pàg. 520. L’autor planteja que l’ajust es podria determinar establint certs factors que queden fora del control de les persones candidates i que tot i així afecten a les seves possibilitats d’admissió. Una vegada confirmat que concorre algun dels factors preestablerts en una persona candidata, se li afegixen cert nombre de punts.

⁵¹ KROLL, J. *et al.*, “Accountable algorithms”, *University of Pennsylvania Law Review*, vol. 165, núm. 3, 2017, pàg. 685.

⁵² *Idem*, pàg. 688.

⁵³ Al llarg del treball, quan ens referim a “directives antidiscriminació europees” estem fent referència a les següents directives: Directiva 2000/43/CE del Consejo, de 29 de junio de 2000, relativa a la aplicació del

la Unió Europea, (d'ara en endavant, TFUE) com a part de les accions per fomentar la igualtat entre homes i dones⁵⁴. Així i tot, en general, el Tribunal de Justícia de la Unió Europea (d'ara en endavant, TJUE) s'ha mostrat reticent a acceptar la introducció d'accions positives. Per exemple, en l'assumpte Kalanke no va permetre la prioritització automàtica del grup desavantatjat, ja que va considerar que una normativa nacional que garantia la preferència absoluta i incondicional de les dones en un nomenament o promoció anava més enllà d'una mesura de foment de la igualtat de tracte⁵⁵.

Interrelacionem seguidament, les decisions automatitzades mitjançant sistemes algorítmics i el concepte “discriminació”.

1.3. La discriminació i els sistemes algorítmics

Discriminar és seleccionar excloent⁵⁶ i entra en l'àmbit de la regulació legal quan suposa donar tracte desigual a una persona o col·lectivitat per motius racials, religiosos, polítics, de sexe, d'edat, etc. Per poder entendre com opera la discriminació impera distingir els conceptes estereotip, prejudici i discriminació. La discriminació és l'acte final, una conducta que ve precedida dels dos elements anteriors: una manera de conèixer la realitat (estereotip) i d'un element axiològic, valoratiu o ideològic (prejudici) sobre aquest estereotip⁵⁷. La jurisprudència del Tribunal Europeu de Drets Humans (d'ara en endavant,

principio de igualdad de trato de las personas independientemente de su origen racial o étnico. DOCE, L 180/22, 19.07.2000; Directiva 2000/78/CE del Consejo, de 27 de noviembre de 2000, relativa al establecimiento de un marco general para la igualdad de trato en el empleo y la ocupación. DOCE, L 303/16, 2.12.2002; Directiva del Consejo 2004/113/CE, de 13 de diciembre de 2004, por la que se aplica el principio de igualdad de trato entre hombres y mujeres al acceso a bienes y servicios y su suministro. DOUE, L 373/37, 21.12.2004; i Directiva 2006/54/CE del Parlamento Europeo y del Consejo, de 5 de julio de 2006, relativa a la aplicación del principio de igualdad de oportunidades e igualdad de trato entre hombres y mujeres en asuntos de empleo y ocupación (refundición). DOUE, L 204/23, 26.7.2006. En relació amb les accions positives, s'hi fa referència en l'article 5 de la Directiva 2000/43/CE, en l'article 7 de la Directiva 2000/78/CE, en l'article 3 de la Directiva 2006/54/CE i a l'article 6 de la Directiva 2004/113/CE.

⁵⁴L'article 157.4 del TFUE disposa que “*Con objeto de garantizar en la práctica la plena igualdad entre hombres y mujeres en la vida laboral, el principio de igualdad de trato no impedirá a ningún Estado miembro mantener o adoptar medidas que ofrezcan ventajas concretas destinadas a facilitar al sexo menos representado el ejercicio de actividades profesionales o a evitar o compensar desventajas en sus carreras profesionales*”.

⁵⁵TJCE, sentència de 17 d'octubre de 1995, C-450/93, Eckhard Kalanke i Freie Hansestadt Bremen, par. 24.

⁵⁶Discriminar. Real Acadèmia Espanyola. 2023. Disponible a: <https://dle.rae.es/discriminar>. Darrera visita el 2 de setembre de 2023.

⁵⁷REY MARTÍNEZ, F., *Derecho Antidiscriminatorio*, op. cit., 2019, pàg. 50.

TEDH), estableix que la discriminació significa tractar de manera diferent, sense una justificació objectiva i raonable, a persones en una situació anàloga o similar de manera rellevant⁵⁸. Afegeix que s'observa quan es tracta diferent a persones o col·lectius en una situació idèntica i quan es tracta igual a persones o grups en situacions diferents⁵⁹.

Per proporcionar una protecció a aquelles persones que es poden veure discriminades, moltes constitucions i tractats de drets humans prohibeixen la discriminació⁶⁰. Binns planteja que l'existència d'animositat o preferència sistemàtica en relació a certs grups socials és el que fa errònia la discriminació, atès que s'arriba a conclusions sobre individus basades en generalitzacions de grups als quals pertanyen⁶¹. En el cas de la presa de decisions mitjançant algoritmes, les conclusions individuals a les que arriba el sistema es basen en generalitzacions fetes des d'una sistematització completament opaca, essent aquesta una de les principals raons del potencial d'aquesta tipologia de discriminacions per evitar les regulacions antidiscriminació existents⁶².

Els algoritmes estan dissenyats per operar basant-se en el fet que una pauta repetida és la més correcta, de manera que la hipòtesi correcta serà aquella que més vegades es repeteixi, sigui correcta o no, ja que la seva capacitat de raonament està fonamentada en la rèplica històrica, reproduint virtualment la realitat i per consegüent, les seves profundes desigualtats socials⁶³. Per aquesta raó, moltes vegades no es tracta de decisions deliberadament discriminatòries, sinó de la traducció en els sistemes algorítmics de la

⁵⁸TEDH (Gran Sala), sentència de 13 de novembre de 2007, 57325/00, D.H. i altres contra la República Txeca, par. 175; TEDH (Secció Quarta), sentència de 12 de desembre de 2006, 13378/05, Burden contra el Regne Unit, par. 60; TEDH (Secció Primera), sentència de 3 d'octubre de 2013, 552/10, I.B. contra Grècia, par. 75.

⁵⁹TEDH (Gran Sala), sentència de 6 d'abril de 2000, 34369/97, Thlimmenos contra Grècia, par. 44. La jurisprudència del TEDH exposa que en una diferència de tracte, les característiques o trets personals són l'element diferencial en els termes comparatius. En la comparació de la persona discriminada respecte a una altra en la mateixa situació, els motius protegits o categories especials són l'element que determina el tracte discriminatori, essent en els trets que caracteritzen a la persona on s'identifica la discriminació. Vegeu, també, per exemple: TEDH (Secció Quarta), sentència de 27 de novembre de 2007, 77782/01, Luczak contra Polònia. En aquest assumpte el tribunal considera que hi ha hagut una vulneració de l'article 14 del Conveni Europeu de Drets Humans, juntament amb l'article 1 del Protocol núm.1, al negar-se-li l'accés a una persona al règim polonès de seguretat social dels agricultors a causa de tenir la nacionalitat francesa.

⁶⁰Per exemple, l'article 7 de la Declaració de Drets Humans de les Nacions Unides, l'article 26 del Pacte Internacional dels Drets Civils i Polítics, l'article 21 de la Carta de Drets Fonamentals de la Unió Europea, l'article 14 del Conveni Europeu de Drets Humans o l'article 14 de la Constitució Espanyola de 1978.

⁶¹BINNS, R., "Fairness in machine learning...", *op. cit.*, 2018, 150-153.

⁶²CRAWFORD, K. i SCHULTZ, J., "Big data and due process: Toward a framework to redress predictive privacy harms", *Boston College Law Review*, vol. 55, núm. 1, 2014, pàg. 99.

⁶³RIVAS VALLEJO, P., "Discriminación algorítmica: detección, prevención y tutela", *XXXI Jornades Catalanes de Dret Social ("Trellball, discriminació i COVID")*, 2021, pàg. 5.

discriminació històrica i estructural a la qual es veuen sotmeses sistemàticament persones que pertanyen a determinats grups en posició de desavantatge.

En aquest sentit, la discriminació estructural esdevé un element central per entendre la discriminació algorítmica, i que no es consideri suposa una limitació a la protecció efectiva dels drets a la igualtat i a la no discriminació⁶⁴. Permet entendre que en tots els casos on es produeix discriminació, aquesta és resultat d'una societat que de manera sistemàtica situa persones que pertanyen a determinats grups socials en una posició de desavantatge⁶⁵. En conseqüència, si no es corregeixen les estructures que situen en un punt de partida desigual a persones pertanyents a grups desavantatjats, no es millorarà en relació a la discriminació⁶⁶. La conceptualització de la discriminació estructural suggereix que la regulació de les formes subtils de discriminació habituals en l'àmbit laboral, per exemple, requereix focalitzar l'atenció en el biaix discriminatori, ja que està influït, habilitat i fins i tot fomentat per les estructures, pràctiques i dinàmiques de les organitzacions i grups dins els quals treballen les persones⁶⁷. S'ha de tenir en compte, així i tot, que l'actuació del dret davant les desigualtats estructurals no només té un àmbit d'efectivitat limitat⁶⁸, sinó que al ser un element que participa en les estructures, té la capacitat de perpetuar-les.

Hi ha una relació clara entre la discriminació estructural i la manca de diversitat de les persones que es dediquen a la intel·ligència artificial. Les dones, des d'un punt de vista marcadament utilitarista, són un recurs poc aprofitat en àmbits emergents de la tecnologia digital, la intel·ligència artificial i les TIC⁶⁹, ja que només representen el 16% de les quasi vuit milions de persones que treballen en aquest sector a Europa⁷⁰. Aquestes dades tenen una incidència directa en les desigualtats i la discriminació de gènere que s'han vist reproduïdes a través del disseny, les aportacions i l'ús de la intel·ligència artificial, com també en generar arxius incomplets i biaixos incorrectes que poden distorsionar la lògica

⁶⁴ AÑÓN ROIG, M. J., "Principio antidiscriminatorio y determinación de la desventaja", *Isonomía: Revista de teoría y filosofía del derecho*, núm. 39, 2013, pàg. 130.

⁶⁵ HAVELKOVÁ, B., "Judicial Scepticism of Discrimination at the ECtHR", a COLLINS, H., KHAITAN, T., (eds.), *Foundations of Indirect Discrimination Law*, Oxford, Hart Publishing, 2018, pàg. 85.

⁶⁶ SORIANO ARANZ, A., "Decisiones automatizadas y discriminación...", *op. cit.*, 2021, pàg. 31.

⁶⁷ MERCAT-BURNS, M., "Systemic discrimination: Rethinking the tools of gender equality", *European Equality Law Review*, núm. 2, 2018, pàg. 5.

⁶⁸ AÑÓN ROIG, M. J., "Principio Antidiscriminatorio...", *op. cit.*, 2013, pàg. 131.

⁶⁹ Acrònim de Tecnologies de la Informació i la Comunicació.

⁷⁰ PARLAMENTO EUROPEO, *Resolución del Parlamento Europeo, de 21 de enero de 2021, sobre la estrategia de la Unión para la igualdad de género*, 2019/2169(INI), Bruselas, 2021, considerando AG.

del sistema i posar en perill la igualtat de gènere en la societat⁷¹. La manca de representació de dones en aquests àmbits educatius manté la segregació de gènere en el mercat laboral i condueix a una manca de varietat de perspectives de gènere. Aquesta manca de veu de les dones i grups minoritaris en el disseny algorítmic té repercussions clares en els biaixos algorítmics⁷².

Tanmateix, malgrat els esforços ben intencionats de les persones professionals de la intel·ligència artificial per incorporar el coneixement sobre el món social als sistemes sociotècnics, generalment no reconeixen la complexitat de la vida social, tendint a posicionar les dades extretes d'éssers humans com a singulars, directes i estables sense tenir en compte com les dinàmiques de poder, la discriminació sistèmica i les desigualtats socials contribueixen al significat de les categories socials; i donant prioritat a allò fonamental per a l'execució del sistema en detriment de la representació en aquest de la complexitat de les desigualtats socials⁷³.

Aquest fet es veu clarament en les bases de dades que s'utilitzen per entrenar sistemes algorítmics. En el llibre “*Atlas of AI*”, Crawford exposa l'exemple de la base de dades⁷⁴ UTKFace -produïda per la Universitat de Tennessee a Knoxville-, que consta de vint mil imatges amb anotacions en relació a l'edat, al gènere, i a l'origen ètnic. En relació a l'edat, les anotacions s'expressen en anys de l'1 al 116, en relació al gènere s'estableix una possibilitat binària, zero és per home i u és per dona, i finalment en relació a l'origen ètnic s'ha de classificar en base a cinc categories: blanc, negre, asiàtic, indi o altres⁷⁵. Òbviament aquests criteris són molt reduccionistes, ja que introdueixen una varietat de lògiques classificatòries que o bé reïfiquen les categories existents per classificar persones o bé en creen de noves⁷⁶. Aquestes tècniques de classificació discriminatòries s'hereten, i traslladen a l'esfera de la intel·ligència artificial la discriminació social estructural. Per

⁷¹ *Ídem, ibidem.*

⁷² GERARDS, J. i XENIDIS, R., *Algorithmic Discrimination in Europe...*, op. cit., 2021, pàg. 61.

⁷³ JOYCE, K. et al., “Toward a Sociology of Artificial Intelligence: A Call for Research on Inequalities and Structural Change”, *Socius: Sociological Research for a Dynamic World*, vol. 7, 2021, pàg. 5.

⁷⁴ Una base de dades es pot construir en recollint dades d'altres fonts o creant-ne de noves. Per exemple, a finals del segle XX, el Departament de Defensa estatunidenc va patrocinar el programa FERET (Facial Recognition Technology) per desenvolupar un sistema automàtic de reconeixement facial. El laboratori d'investigació de l'exèrcit dels Estats Units va liderar un projecte tècnic creant fotografies de més d'un miler de persones en diferents posicions per assolir un total de 14.126 imatges, esdevenint FERET un punt de referència estandarditzat (una eina compartida de mesura per comparar perspectives per detectar cares).

⁷⁵ CRAWFORD, K., *Atlas of AI*, New Haven and London, Yale University Press, 2021, pàg. 144-145.

⁷⁶ STRAK, L., “Facial Recognition is the Plutonium of AI”, *XRDS: Crossroads, The ACM Magazine for Students*, vol. 25, núm. 3, 2019, pàg. 55.

exemple, ImageNet -una base de dades d'imatges- hereta la tipologia de classificació de WordNet -una base de dades de lèxic de la dècada de 1980-, que al seu torn és hereva de moltes fonts entre elles el Brown Corpus d'un milió de paraules publicat el 1961. Sorgeix així, una genealogia de sistemes d'entrenament que han heretat les lògiques apreses de mostres anteriors i donen lloc a les subsegüents. Aquest funcionament de construcció sobre aquelles bases de dades ja existents, crea els límits que governen com la intel·ligència artificial opera i com entén el món⁷⁷.

Seguidament exposem més concretament com opera la discriminació algorítmica i els elements que la caracteritzen.

2. Causes de discriminació en les decisions automatitzades mitjançant sistemes algorítmics

Tècnicament s'entén per algoritme un conjunt concret d'instruccions que serveixen per executar una tasca. Es pot categoritzar la funció i el poder dels algoritmes en quatre grans categories: a) la classificació -categorització de la informació basada en les característiques que presenta en diferents aparats-; b) la priorització -denotació d'èmfasi i rang d'una informació o resultats particulars a expenses d'altres basada en uns criteris prèviament establerts-; c) l'associació -determinació de correlacions entre entitats-; i d) el filtratge -inclusió o exclusió d'informació basada en criteris prèviament determinats⁷⁸. Bàsicament la tasca de l'anàlisi massiva de dades consisteix en localitzar relacions estadístiques entre un conjunt de dades, de manera que s'automatitza el procés de descobriment de patrons útils que operen regularment de la mateixa manera. El conjunt de relacions que es descobreix s'anomena model o model predictiu. Aquest model, es podrà aplicar a qualsevol conjunt de dades, ja que haurà après quins atributs i activitats són rellevants i així poder treure conclusions de les relacions que es creen entre elles. Per crear un model és necessari, per una banda, l'etiquetatge de totes les dades rellevants en categories mútuament excloents, i per altra banda, definir les variables, que determinen

⁷⁷ CRAWFORD, K., *Atlas of AI, op. cit.*, 2021, pàg. 98.

⁷⁸ DIAKOPOULOS, N., "Algorithmic accountability: Journalistic investigation of computational power structures", *Digital Journalism*, vol. 3, núm. 3, 2015, pàg. 400-402.

quins són els objectius de l'anàlisi de dades. L'etiquetatge i la concreció de variables són processos subjectius que duen a terme aquelles persones encarregades de l'anàlisi de dades⁷⁹. En el sector privat l'ús d'algoritmes pot servir, per exemple, per avaluar la capacitat de crèdit de la clientela en el sector bancari⁸⁰, en l'assistència sanitària⁸¹ o en els recursos humans⁸². En el sector públic, pot servir, per exemple, en els departaments de policia⁸³, en la gestió del sistema de justícia penal⁸⁴ o per l'anàlisi de dades massives poblacionals agregades i anonimitzades, amb la finalitat d'obtenir informació en temps real sobre el comportament de la multitud, predicció d'àrees de risc o adaptació d'intervencions de polítiques públiques⁸⁵.

Sáez Lara identifica tres elements que caracteritzen la discriminació algorítmica: la invisibilitat, que es projecta en la creença en l'objectivitat de les dades determinant una acceptació acrítica d'aquests sistemes, sense anàlisi ni avaluació i en l'absència de responsabilitat humana sobre els resultats esbiaixats o discriminatoris donada l'opacitat de la participació humana en la decisió⁸⁶; la intensitat, ja que donat l'elevat volum, velocitat i varietat de dades usades, qualsevol biaix introduït en el sistema magnificarà i multiplicarà els seus efectes i; la complexitat tècnica, ja que els sistemes algorítmics no

⁷⁹BAROCAS, S. i SELBST, A. D., "Big Data's Disparate Impact", *California Law Review*, vol. 104, 2016, pàg. 678.

⁸⁰Vegeu, en aquest sentit, per exemple: ZUIDERVEEN, F., *Discrimination, artificial intelligence, and algorithmic decision-making*, Council of Europe, Strasbourg, 2018, pàg. 7.

⁸¹Vegeu, en aquest sentit, per exemple: CHEN, M. *et al.*, "Disease prediction by Machine learning over big data from healthcare communities", *IEEE Access*, vol. 5, 2017, pàg. 8869-8879.

⁸²Vegeu, en aquest sentit, per exemple: KIRIMI, J.M. i MOTURI, C. A., "Application of data mining classification in employee performance prediction", *International Journal of Computer Applications*, vol. 146, núm. 7, 2016, pàg. 28-35.

⁸³Vegeu, en aquest sentit, per exemple: SELBST, A. D., "Disparate Impact in Big Data Policing", *Georgia Law Review*, vol. 52, núm. 1, 2017, pàg. 109-195.

⁸⁴Vegeu, en aquest sentit, per exemple: CABALLÉ-PÉREZ, M. *et al.*, "El quebrantamiento de las órdenes de protección en violencia de género: análisis de los indicadores de riesgo mediante el formulario vpr4.0", *Anuario de psicología jurídica*, núm. 30, 2020, pàg. 63-72.

⁸⁵OBSERVATORI D'ÈTICA EN INTEL·LIGÈNCIA ARTIFICIAL DE CATALUNYA (OEIAC), *Intel·ligència artificial, ètica i societat...*, *op. cit.*, 2021, pàg. 32. Per alguns exemples dels usos d'algoritmes en els estats membres de la UE, vegeu: ALLEN, R. i MASTERS, D., *Regulating for an equal AI: A new role for equality bodies. Meeting the new challenges to equality and non-discrimination from increased digitisation and the use of Artificial Intelligence*, European Network of Equality Bodies (EQUINET), Brussels, 2020, pàg. 34 i 35.

⁸⁶En aquest sentit, vegeu també: TORRALBA, F., *L'ètica algorítmica*, Barcelona, Edicions 62, 2022, pàg. 79. L'autor exposa que malgrat es pensi que les dades massives ens alliberen de la part d'incertesa o error en què poden caure els éssers humans en emetre un judici o donar un consell o una recomanació, és probable que produeixin una sèrie d'errors i injustícies. Afirmar que tot aquest "progrés" s'institueix en contra de la idea de política i democràcia, sobretot perquè, generalment, els creadors dels algoritmes no estan subjectes a cap responsabilitat pública ni retiment de comptes. Exposa que mentre que la intolerància s'emascara de tolerància en la societat controladora, en la societat algorítmica l'opacitat s'emascara de transparència però la gestió de les dades massives està lluny de ser transparent, ja que reforça la dinàmica de privatització del poder i cultiu de la cultura del secret.

només es basen en una correlació o unió de dades sinó que també, en la interpretació de les dades, sorgint així biaixos en les dades però també en la forma en què l'algoritme treballa i la seva facultat de predicció derivada de biaixos presents en la societat⁸⁷.

Interessa exposar quin és el funcionament de la presa de decisions automatitzades mitjançant sistemes algorítmics per discriminar certs col·lectius o individus en base a diverses condicions (sexe, origen racial, religió, situació socioeconòmica, etc.), atès que s'ha posat de relleu que els processos de desenvolupament i posada en funcionament dels algoritmes ofereixen una àmplia sèrie de possibilitats per a la incorporació de biaixos que perjudiquen els membres de grups desavantatjats⁸⁸. Per exposar el funcionament de la presa de decisions automatitzades mitjançant sistemes algorítmics és necessari fer referència a un estudi de Barocas i Selbst⁸⁹, on es distingeixen les diverses formes en les quals la presa de decisions automatitzades pot conduir a la discriminació. També n'hem afegit d'altres que malgrat no estar plantejades en l'estudi anteriorment referenciat, també considerem rellevants.

Hem dividit les formes de discriminació en aquelles que succeeixen en base a mesures tècniques (2.1), és a dir, que ocorren a causa del funcionament tècnic del sistema algorítmic de presa de decisió (malgrat que les persones hagin intervingut indirectament, com per exemple, introduint les dades i creant les bases de dades, creant les categories, les classificacions, etc.); i aquelles que succeeixen en base a apreciacions humanes (2.2), és a dir, aquelles que són causades directament i intencionalment per les persones que intervenen o exploten les decisions generades pel sistema algorítmic.

2.1. Discriminació en base a mesures tècniques

Podem classificar la discriminació segons si és conseqüència d'errors i biaixos en les bases de dades d'entrenament dels algoritmes (2.1.1), de la selecció i pes atorgat a les

⁸⁷SÁEZ LARA, C., "El algoritmo como protagonista de la relación laboral. Un análisis desde la perspectiva de la prohibición de discriminación", *Temas Laborales: Revista andaluza de Trabajo y bienestar social*, núm. 155, 2020, pàg. 45-46.

⁸⁸BAROCAS, S. i SELBST, A. D., "Big Data's Disparate...", *op. cit.*, 2016, pàg. 677.

⁸⁹*Idem*, pàg. 671.

variables (2.1.2), de la selecció de les característiques de classificació de les dades (2.1.3), de l'obtenció de dades personals mitjançant correlacions amb altres dades (2.1.4), de l'assumpció o l'associació de les dades amb grups socials discriminats (2.1.5) o de la perpetuació d'estereotips en els quals es basen les estructures socials de discriminació (2.1.6).

2.1.1. Errors i biaixos en les bases de dades d'entrenament de l'algoritme

La discriminació pot venir derivada dels errors i biaixos continguts en les bases de dades emprades en el desenvolupament dels sistemes de presa de decisions automatitzades. Pot ocórrer en la construcció de la base de dades (2.1.1.1) o després de la construcció (2.1.1.2)⁹⁰.

2.1.1.1. En la construcció de la base de dades

Freqüentment, les bases de dades s'han construït sense que la persona les dades de la qual s'empren sigui conscient de l'abast que aquestes poden tenir. Les bases de dades NIST⁹¹, per exemple, que estan construïdes en base a fotografies policials de persones detingudes (en anglès mugshots) sobre les quals no s'ha demanat el consentiment, i que serveixen per desenvolupar i provar sistemes automatitzats d'identificació⁹². També, moltes bases de dades que estan disponibles gratuïtament per al desenvolupament de la intel·ligència artificial plenes d'imatges de persones duent a terme accions quotidianes com seure,

⁹⁰ SURESH, H. i GUTTAG, J., "Framework for Understanding Unintended Sources of Harm Throughout the Machine Learning Life Cycle", *EAAMO 2021: ACM Conference on Equity and Access in Algorithms, Mechanisms, and Optimization*, 2021, pàg. 4-6. Els autors apunten que les persones sovint parlen de biaix en els sistemes d'aprenentatge automàtic sense arribar a distingir entre efectes del biaix que poden venir de diferents fases del procés d'aprenentatge automàtic (incloent el biaix històric que es previ al desenvolupament) i del biaix de desplegament que està relacionat amb com el sistema s'usa i que és, consegüentment, posterior. En base a aquest apunt hem dividit els subapartats.

⁹¹ Es tracta de bases de dades científiques públiques del National Institute of Standards and Technology (NIST) estatunidenc.

⁹² CRAWFORD, K., *Atlas of AI, op. cit.*, 2021, pàg. 92.

saludar, riure o plorar, han estat fetes sense ni el coneixement ni el consentiment de les persones implicades i recollides de fonts en línia com Google o Youtube⁹³. A més, no cal que les persones no en tinguin coneixement o que l'accés sigui gratuït, sinó que les companyies que desenvolupen la IA segueixen amb intenció d'arribar a àrees de la vida humana que es consideraven fora dels límits o que suposaven massa cost⁹⁴. La investigació sobre el pacte entre el Britain's Royal Free National Health Service Foundation Trust i l'empresa DeepMind (subsidiària de Google) per compartir informes d'1'6 milions de pacients va revelar que s'havia violat la legislació de protecció de dades, ja que no s'havia informat de manera suficient les persones⁹⁵ sobre les conseqüències que comportaria la cessió de les seves dades.

També, les dades que s'inclouen poden venir de diverses fonts amb la intenció de reunir el màxim de dades possibles desvinculant-les del context del que provenen amb l'objectiu de concebre totes les dades com a reutilitzables mentre hi hagi suficient contingut per entrenar els models en base a l'assumpció que totes les dades són intercanviables⁹⁶. El Penn Treebank Project, per exemple, una base de dades de text creada entre 1989 i 1992 per un equip de lingüistes i informàtics de la Universitat de Pennsilvània amb el propòsit d'entrenar sistemes de processament de llenguatge, incloïa entre les seves fonts sumaris del Departament d'Energia estatunidenc, articles de notícies del Daw Jones o informes del servei federal estatunidenc de notícies sobre “activitat terrorista” a Sud Amèrica⁹⁷. Una altra base de dades de text, va venir de les investigacions de frau de la Corporació Enron després de declarar-se en fallida econòmica, i incloïa entre les seves fonts correus electrònics de 158 persones treballadores confiscats per la Comissió Federal de Regulació de l'Energia⁹⁸, dels quals s'havia fet públic el contingut, creant-se així, una base de dades textual en base a mig milió d'intercanvis en discurs quotidià. Al concebre's aquestes col·leccions de text com a conjunts neutres de llenguatge, s'establia una equivalència entre paraules en un manual tècnic i aquelles emprades amb col·legues via correu electrònic. Malgrat aquesta equivalència, el llenguatge no és neutre i representa un temps,

⁹³ *Ídem*, pàg. 119-120.

⁹⁴ *Ídem, ibídem*.

⁹⁵ *Ídem, ibídem*.

⁹⁶ *Ídem*, pàg. 102-103.

⁹⁷ MARCUS, M.P., SANTORINI, B. i MARCINKIEWICZ, M.A., “Building a Large Annotated Corpus of English: The Penn Treebank”, *Computational Linguistics, Special Issue on Using Large Corpora: II*, vol. 19, núm. 2, 1993, pàg. 313-330.

⁹⁸ KLIMT, B. i YANG, Y., “Introducing the Enron Corpus”. *First Conference on Email and Anti-Spam (CEAS)*, 2004, pàg. 1-2.

un lloc, una cultura i unes polítiques, fent que els llenguatges que tenen menys dades disponibles no serveixin des d'aquesta perspectiva i se'ls deixi enrere⁹⁹, o estiguin menys representats. Aquesta descontextualització també té una incidència en la presa de decisions. Si bé es cert que el coneixement humà contextual pot tenir clarament efectes negatius (portant els biaixos i la discriminació a la presa de decisions), també duu a terme una funció important, ja que suposa que les decisions preses per éssers humans, majoritàriament, no produiran explicacions particularment aleatòries, és a dir, explicacions fora de context que veurem segurament injustificades (com relacionar decisions respecte préstecs bancaris amb l'elecció de mòbil o el color dels mitjons), evitant de la mateixa manera certs tipus d'error extrem¹⁰⁰.

2.1.1.2. Després de la construcció de la base de dades

No només la construcció de la base de dades, sinó que l'etiquetatge de dades també pot presentar conflictes. Per una banda, inclús quan els etiquetatges no creen controvèrsia, poden presentar problemes ja que les persones analistes han de realitzar eleccions difícils per decidir quins etiquetatges són més adequats per a cada exemple (i cada exemple pot presentar alguns criteris, però no tots, per ser inclòs en cada etiquetatge). Per altra banda, pot passar que l'etiquetatge no sigui suficientment precís per capturar diferències significatives entre els casos. Aquestes situacions comporten que les persones analistes duguin a terme judicis de valor amb rellevància en la presa de decisions posterior¹⁰¹. Un cas representatiu pot ser el d'ImageNet, una base de dades que funciona amb la sistemàtica de classificació d'imatges sota nou categories (que després es desglossen en subcategories): planta, formació geològica, objecte natural, esport, artefacte, fong, persona, animal i divers. Malgrat conceptes com "poma" poden ser fàcils de catalogar, n'hi ha d'altres més complicats. Quan ens referim a la categoria "persona", s'integren etiquetes com "persona sense habilitats" o "mulatto" que integren al seu torn imatges on s'hi veuen moltes assumpcions i estereotips incloent raça, gènere, edat i habilitat. No hi

⁹⁹ BAKER, J., DENG, L., GLASS, J., KHUDANPUR, S., LEE, C. H., MORGAN, N. i O'SHUGHNESSY, D., "Research Developments and Directions in Speech Recognition and Understanding. Part 1", *IEEE Signal Processing Magazine*, vol. 26, núm. 3, 2009, pàg. 75-80.

¹⁰⁰ KAMINSKI, M.E., "Binary Governance...", *op. cit.*, 2019, pàg. 1547.

¹⁰¹ BAROCAS, S. i SELBST, A. D., "Big Data's Disparate...", *op. cit.*, 2016, pàg. 681.

ha categories neutres a ImageNet, ja que la selecció d'imatges sempre interactua amb el significat de les paraules. La creació de la classificació d'ImageNet, prové de treballadors als quals se'ls dona la tasca de classificar en les categories donades cinquanta imatges (extretes, per exemple, de navegadors web com Google) per minut. D'aquesta manera, en la base de dades es poden trobar estereotips, errors i absurditats com que una persona en una tovallola a la platja sigui "cleptomana" o una persona adolescent amb una dessuadora d'esports sigui classificada com "perdedora". Malgrat s'intentin suprimir o matisar els pitjors exemples, la perspectiva està fonamentalment construïda en una separació entre les dades i les persones, i els llocs i els contextos d'on provenen. Aquesta és una visió del món tècnica, que pretén fusionar una forma d'objectivitat d'allò que són diversos materials culturals complexos¹⁰².

També pot passar que les dades amb les quals es vol prendre la decisió no siguin representatives¹⁰³. Per exemple, si les dades incorporades en una base de dades sobre delictes no representen de manera acurada l'activitat criminal d'una àrea concreta, ja que aquelles incorporades es limiten a aquelles activitats que les persones decideixen denunciar o allò que els/les policies directament observen (es podria donar el cas que la policia detingués més a persones d'origen immigrant, de manera que centrant la seva atenció en certs grups ètnics i certs barris, és probable que els registres policials sobrerrepresentin sistemàticament a aquests grups i barris, fent que el procés de recollida de dades estigui esbiaixat¹⁰⁴). Així, la llacuna que existeix entre els delictes comesos i els denunciats no és aleatòria sinó sistèmica, lligada per exemple, a la tipologia de delicte -violència de gènere o robatoris-, les característiques, la disponibilitat i accessibilitat de les dades -com per exemple les xarxes socials comparades amb totes les comunicacions-, quines situacions denunciades constitueixen un delicte, com es categoritza i classifica un delicte, o quins requisits existeixen per denunciar una situació¹⁰⁵. També pot succeir la infra o supra representació de persones pertanyents a determinats col·lectius en les dades a causa que no totes les dades són creades o recollides d'igual manera. Pot passar que la qualitat i la representativitat dels registres variï de manera correlacionada amb característiques determinades de les persones –per exemple, les institucions poden

¹⁰² CRAWFORD, K., *Atlas of AI, op. cit.*, 2021, pàg. 138-147.

¹⁰³ HOUSE OF LORDS SELECT COMMITTEE ON ARTIFICIAL INTELLIGENCE, *AI in the UK: Ready, willing and able*, London, 16 April 2018, pàg. 41-44.

¹⁰⁴ LUM, K. i WILLIAM, I., "To predict and serve?", *Significance*, vol.13, núm. 5, 2016, pàg. 15.

¹⁰⁵ BENNET MOSES, L. i CHAN, J., "Algorithmic prediction in policing: assumptions, evaluation, and accountability", *Policing and Society*, vol. 28, núm. 7, 2018, pàg. 810.

mantenir sistemàticament registres menys precisos i complets per certs col·lectius de persones-, fracassant en representar diferents grups en les proporcions acurades¹⁰⁶.

En errors en les bases de dades poden caure de manera més freqüent persones pertanyents a grups que s'han vist en desavantatge històricament perquè a causa de la pobresa, la geografia o la manera de viure participen menys en l'economia formal i les seves activitats generadores de dades, tenen un accés desigual a la tecnologia¹⁰⁷, són consumidores menys rendibles i per tant, menys interessants com a subjectes d'observació¹⁰⁸. D'elles, es disposa d'un volum de dades anormalment baix (estan menys "datificades" que la població general, per consegüent resulten particularment difícils d'aplicar les tècniques de la IA), considerant-se que pateixen discriminació econòmica¹⁰⁹. Un exemple d'aquesta situació es dona en una aplicació de Boston "Street Bump" que s'aprofita dels acceleròmetres dels telèfons mòbils per detectar quan els/les conductors/es passen per sobre de sots al paviment. Tot i el pragmatisme de l'aplicació, tota la informació que l'ajuntament rebí d'aquesta aplicació estarà esbiaixada a causa de la desigual distribució de telèfons mòbils amb possibilitat de tenir aquesta aplicació entre la població de la ciutat. Les diferències sistemàtiques en relació a la propietat dels telèfons suposarà que les comunitats més pobres on es congreguen persones més vulnerables no generaran informació sobre problemes en el paviment, de manera que si la ciutat ha de destinar recursos econòmics a solucionar problemes segons l'aplicació, s'estarà discriminant contra aquells qui tenen menys recursos econòmics¹¹⁰.

També poden trobar-se errors que condueixin a resultats discriminatoris a causa de la sobrealimentació de l'algoritme en vistes del seu entrenament amb bases de dades on hi ha major representació de certs grups (per exemple, persones d'origen europeu amb determinats trets)¹¹¹, o amb dades relatives a trets que no tenen cap relació amb la finalitat

¹⁰⁶BAROCAS, S. i SELBST, A. D., "Big Data's Disparate...", *op. cit.*, 2016, pàg. 684.

¹⁰⁷ En relació amb l'ètica, la justícia i l'accés desigual a la tecnologia, vegeu: TORRALBA, F., *L'ètica algorítmica, op.cit.*, 2022, pàg. 12. L'autor exposa que la bretxa tecnològica és visible en els grups vulnerables econòmicament i socialment, qüestió que no és irrellevant a l'hora d'abordar l'ètica de la tecnologia i, en particular, l'ètica algorítmica, perquè el fonament principal de l'ètica és la justícia, i la justícia consisteix a garantir una distribució equitativa dels recursos, ja siguin naturals o artificials.

¹⁰⁸LERMAN, J., "Big Data and its exclusions", *Stanford Law Review*, vol. 66, 2013, pàg. 57.

¹⁰⁹En aquest sentit, a la doctrina espanyola, vegeu: SANCHO LÓPEZ, M., "Estrategias legales para garantizar los derechos fundamentales frente a los desafíos del Big Data", *Revista General de Derecho Administrativo*, núm. 50, 2019, pàg. 4. L'autora fa referència a una societat de classes digital.

¹¹⁰CRAWFORD, K., "Think again: Big Data", *Foreign Policy Review*, 10 de maig de 2013. Disponible a: <https://foreignpolicy.com/2013/05/10/think-again-big-data/>. Darrera visita el 2 de setembre de 2023.

¹¹¹ HIGH LEVEL GROUP ON NON-DISCRIMINATION, EQUALITY, AND DIVERSITY. SUBGROUP ON EQUALITY DATA, *Guidelines on improving the collection and use of equality data*, Brussels, 2018,

per la qual seran aplicats els seus resultats, però que el mecanisme d'anàlisi de l'algoritme inclourà de la mateixa manera. Un exemple de sobrealimentació el podem trobar en els algoritmes predictius emprats per ocupar un lloc de feina (la fase d'entrenament de l'aprenentatge automàtic s'ha identificat com la principal font de biaixos en els processos de contractació¹¹²) que es basen en les característiques de les persones candidates en comptes d'en les seves aptituds demostrades. Els algoritmes entrenats per obtenir un resultat en base a una combinació de dades, gestionades quasi en la seva totalitat de manera digitalitzada, poden tendir a una expansió innecessària, més enllà de les característiques que resulten necessàries per cobrir el lloc de treball, i que inclús potser no van ser proporcionades per les persones candidates en qüestió¹¹³.

Finalment, pot passar, simplement, que les dades d'entrenament no introdueixin biaixos sinó que reproduïxin aquells ja existents. Un exemple d'aquesta situació la representa el cas de L'Escola de Medicina de l'Hospital de Sant George¹¹⁴. Aquesta escola va rebre més candidatures de les que podia admetre i va desenvolupar un programa informàtic per determinar quines candidatures s'havien d'admetre. Les dades d'entrenament per l'admissió eren fitxers d'anys previs, quan les persones decidien qui entrava a la facultat de medicina, i mostraven quines característiques es correlacionaven amb l'objectiu final (admissió a la facultat de medicina). El sistema discriminava envers dones i persones d'origen immigrant, ja que el criteri de les persones que feien la selecció anteriorment estava esbiaixat envers aquests col·lectius¹¹⁵.

2.1.2. Selecció i pes atorgat a les variables

pàg. 21. Es fa referència en aquestes guies a quan la base de dades és suficientment àmplia però esbiaixada perquè les persones que formen part d'una població concreta que pateix més discriminació formen part en més o menys mesura de l'enquesta, recomanant que es dugui a terme una recollida de dades inclusiva.

¹¹²PÁEZ, A., "Negligent Algorithmic Discrimination", *Law and Contemporary Problems*, vol. 84, núm. 3, 2021, pàg. 19.

¹¹³RIVAS VALLEJO, P., "Discriminación algorítmica: detección, prevención y tutela", *op. cit.*, pàg. 21 i 22.

¹¹⁴ En el sistema d'accés a la universitat estatunidenc, cada universitat estableix els seus propis estàndards o normatives d'admissió.

¹¹⁵LOWRY, S. i MACPHERSON, G., "A blot on the profession", *British Medical Journal*, vol. 296, núm. 6623, 1988, pàg. 657-658.

A través de la selecció i pes atorgats a les variables emprades pel sistema es poden crear algoritmes que reproduïxin les estructures tradicionals de discriminació de grups de persones desavantatjats. Per tant, la forma en la que es dona prioritat a unes variables sobre d'altres pot condicionar i esbiaixar els resultats¹¹⁶. En moltes ocasions, els algoritmes discriminen en base a criteris de mesurament aparentment neutres però que perjudiquen en major mesura als membres de grups desavantatjats¹¹⁷.

Un exemple senzill de la selecció de les variables es el següent: l'algoritme empra dades d'entrenament (milions de missatges de correu electrònic, dels quals un 60% ha estat etiquetat com a correu no desitjat per persones en el passat, i el 40% restant se suposen correus desitjats). Els sistema algorítmic pot trobar correlacions entre el contingut dels missatges, aprenent, per exemple, a reconèixer missatges de correu no desitjat en base a les característiques del correu (variables) com la seva llargada, les paraules incloses en el missatge o errors d'ortografia (p. e. “pastilla màgica per la pèrdua de pes” o “has guanyat un milió d'euros”). En base a aquestes característiques (variables) l'algoritme pot diferenciar entre correu desitjat i no desitjat amb nous missatges (que no formen ja part de la base de dades)¹¹⁸. Aquesta tasca pot esdevenir més complexa quan, per exemple, s'empra aquest sistema per buscar una persona per realitzar una feina i l'objectiu és que sigui “bona” a la feina. Hi ha molts factors que fan que a una persona se la consideri “bona” treballadora. Imaginem llavors, que un sistema d'intel·ligència artificial té en compte la puntualitat com a variable. Aquest fet pot perjudicar persones amb menor poder adquisitiu atès que és més probable que visquin als afores de la ciutat, on hi ha més possibilitat de trobar embussos de trànsit o a problemes amb el transport públic que puguin afectar la seva puntualitat¹¹⁹. D'aquesta manera, s'està tenint en compte una variable que pot, de manera indirecta, aportar informació perjudicial per certs grups de persones vulnerables.

¹¹⁶SORIANO ARNANZ, A., “Decisiones automatizadas y discriminación...”, *op. cit.*, 2021, pàg. 11.

¹¹⁷ CECCHI-DIMEGLIO, P., “How gender bias corrupts performance reviews, and what to do about it”, *Harvard Business Review*, 12 d'abril de 2017. Disponible a: <https://hbr.org/2017/04/how-gender-bias-corrupts-performance-reviews-and-what-to-do-about-it>. Darrera visita el 2 de setembre de 2023.

¹¹⁸ GERARDS, J. i ZUIDERVEEN, F., “Protected Grounds and the System of non-discrimination law in the context of algorithmic decision-making and artificial intelligence”, *Colorado Technology Law Journal*, vol. 20, núm. 1, 2022, pàg. 6.

¹¹⁹ZUIDERVEEN, F., *Discrimination, artificial intelligence...*, *op. cit.*, 2018, pàg. 17.

2.1.3. Selecció de característiques de classificació de dades

Les dades són representacions simples d'un objecte o fenomen infinitament més específic i complex. Es pretén tenir una concepció completa a través de les dades i són les organitzacions i persones que empren el sistema de presa de decisions mitjançant algorismes, les qui determinen quines d'aquestes dades seran rellevants i es tindran en compte en l'anàlisi¹²⁰. Com plantegen Lobacheva i Kashtanova, els algorismes que operen en l'aprenentatge automàtic es desenvolupen emprant dades seleccionades per certs participants privilegiats, en l'interès de persones consumidores semblants a ells, que formen un subgrup de població no representatiu¹²¹. En resum, quan es programa, es pren partit¹²².

Aquestes representacions simplistes manquen la capacitat de captar tots els factors ambientals i mai poden reflectir, com pretenen, la complexitat de les persones¹²³. És problemàtic perquè les representacions simples es donen més en detriment de persones pertanyents a grups vulnerables, a causa dels raonaments estadístics, inherentment generalistes però no universals, fent que certes persones siguin automàticament encaixades per inferències estadísticament sòlides que, tanmateix, són inexactes¹²⁴. Per exemple, seguint la il·lustració que hem plantejat en l'apartat anterior, al concretar la perspectiva d'un/a bon/a treballador/a en si ha estudiat en una universitat prestigiosa¹²⁵ es discrimina envers certs grups de població que rarament tenen accés a aquest tipus d'universitats. Es tracta d'un criteri molt simplista per inferir sobre la capacitat laboral de

¹²⁰BAROCAS, S. i SELBST, A. D., "Big Data's Disparate...", *op. cit.*, 2016, pàg. 688.

¹²¹LOBACHEVA, A. i KASHTANOVA, E., "Social Discrimination in the epoch of artificial intelligence", *Wisdom, Special Issue*, vol. 1, núm. 2, 2022, pàg. 102.

¹²²O'NEIL, C., *Weapons of Math Destruction. How big data increases inequality and threatens democracy*, New York, Broadway Books, 2017, pàg. 20; TORRALBA, F., *L'ètica algorítmica*, *op. cit.*, 2022, pàg. 31. L'autor planteja que una part important de la feina en ètica algorítmica rau a identificar els biaixos implícits i preguntar-se si és legítim automatitzar-los i reproduir-los en les màquines intel·ligents que utilitzem.

¹²³CALDERS, T. i ŽLIOBAITĚ, I., "Why Unbiased Computational Processes Can Lead to Discriminative Decision Procedures", *Discrimination and privacy in the information society*, vol. 3, núm. 20, 2013, pàg. 47; COTINO HUESO, L., "Discriminación, sesgos e igualdad de la inteligencia artificial en el sector público", a GAMERO CASADO, E., (dir.) i PÉREZ GUERRERO, F. L., (coord.), *Inteligencia artificial y sector público. Retos, límites y medios*, Valencia, Tirant lo Blanch, 2023, pàg. 257-351.

¹²⁴BAROCAS, S., i SELBST, A. D., "Big Data's Disparate...", *op. cit.*, 2016, pàg. 688.

¹²⁵Ens referim en aquest exemple al sistema universitari estatunidenc, interpretant universitats prestigioses com a aquelles famoses i cares.

les persones candidates, que a causa de ser tan reduccionista fa que els seus efectes siguin discriminatoris¹²⁶.

De la mateixa manera, en l'àmbit policial i forense estatunidenc, s'empra COMPAS¹²⁷, un algoritme que treballa amb les respostes de persones que han comés un delictes prèviament a un qüestionari del qual deriva la seva probabilitat de reincidència. No fa ús de l'origen racial o el color de pell com a dada, però s'ha afirmat que està esbiaixat envers les persones negres. Si bé prediu la reincidència correctament el 61% de les vegades, les persones negres es veuen etiquetades gairebé el doble de vegades com a risc alt que les persones blanques, sense efectivament reincidir posteriorment; fent l'error oposat entre les persones blanques, és a dir, és molt més probable que se les etiqueti com a risc baix i després sí que reincideixin¹²⁸. Per consegüent, l'algoritme pot estar molt esbiaixat, basat en prejudicis socials i en fonaments racistes i inclús discriminatoris¹²⁹. És un bon exemple del perill de judicis de culpabilitat -aspecte, per definició, individual-, sobre la base de correlacions probabilístiques que consideren, en definitiva, exclusivament l'encaix de la persona en un determinat perfil de risc¹³⁰.

No només això, sinó que la diferenciació es pot donar a través de l'ús de dades irrellevants. Per exemple, algunes companyies d'assegurances dels Països Baixos cobraven més diners per l'assegurança del cotxe si les persones consumidores vivien en apartaments amb cert tipus de numeració (incloent lletres, per exemple 4A o 3B). Les companyies van trobar correlacions en dades històriques entre viure en una adreça d'aquesta tipologia i tenir un accident de trànsit. Tot i així, l'explicació sobre la rellevància d'aquest criteri no sembla lògica tenint en compte la manca de nexes causal entre el número (i lletra) de l'adreça amb les capacitats de conducció¹³¹.

¹²⁶ZUIDERVEEN, F., *Discrimination, artificial intelligence...*, op. cit., 2018, pàg. 20.

¹²⁷Acrònim de "Correctional Offender Management Profiling for Alternative Sanctions". L'ús d'aquest sistema va donar lloc al cas Loomis. Vegeu: State v. Loomis, 881, N.W.2d 749, 7532 (Wis. 2016).

¹²⁸ZUIDERVEEN, F., *Discrimination, artificial intelligence...*, op. cit., 2018, pàg. 24 i 25. L'autor referencia un estudi d'ANGWIN, J. et al., "Machine bias: There's software used across the country to predict future criminals. And it's biased against blacks", publicat a ProPublica el 23 de maig de 2016.

¹²⁹RIVAS VALLEJO, P., "Discriminación algorítmica: detección, prevención y tutela", op. cit., 2021, pàg. 11.

¹³⁰En aquest sentit, en la doctrina espanyola, vegeu: NIEVA FENOLL, J., *Inteligencia artificial y proceso judicial*, Madrid, Marcial Pons, 2018, pàg. 153-154.

¹³¹ GERARDS, J. i ZUIDERVEEN, F., "Protected Grounds and the System...", op. cit., 2022, pàg. 14 i 15.

2.1.4. Obtenció de dades personals mitjançant correlacions amb altres dades

Una altra manera de discriminar és a través de l'ús de dades en aparença neutres que revelen atributs personals. Aquestes dades neutres es poden anomenar indicadors¹³², i s'entenen per part de les persones que programen els sistemes algorítmics com a elements que estan a prop i per tant poden servir com a aproximació a altres elements als quals no tenen accés directe. En un model de predicció algorítmica, es fixa un objectiu (per exemple, evitar la reincidència o la comissió de delictes durant un permís penitenciari), i el model analitza les dades de tots els permisos concedits per intentar trobar quins factors diferencials existeixen entre els permisos que acaben amb la comissió d'un delicte i aquells que no. S'aïllen així uns factors que es prenen en consideració per atorgar o denegar el permís, tot i que ningú els fixa de forma expressa i poden canviar en funció de les dades que s'analitzin¹³³.

L'ús d'aquests indicadors pot portar a la discriminació quan la pertinença a grups vulnerables resulta estar codificada en altres dades¹³⁴. Un exemple per entendre aquesta forma de discriminar és amb la relació entre els codis postals i el pagament dels préstecs bancaris. Un sistema d'intel·ligència artificial emprat per predir quines persones deutores tindran problemes per pagar els préstecs (objectiu final fixat), que no conté informació sobre l'ètnia, aprèn que les persones que pertanyen a cert codi postal tendeixen a tenir més problemes per pagar els préstecs (factor diferencial entre aquelles persones que paguen el préstec i aquelles que no) i ho utilitza com a criteri de predicció per a futures persones susceptibles d'impagaments (fixant de manera autònoma aquest criteri). Tot i semblar un criteri neutral, pot existir una relació directa entre el codi postal i l'ètnia, que suposi que el banc denegui préstecs a persones amb aquest codi postal, perjudicant, a la pràctica, persones amb un origen ètnic determinat¹³⁵. Com expressa Añón Roig, quan ens referim a la intel·ligència artificial, el principi de causalitat es dilueix en pures correlacions, reconeixement de patrons i multiplicació exponencial de les interaccions de dades, incorporades en volums massius amb models molt abstractes que s'entrenen a si mateixos, donant lloc a sistemes complexos que generen opacitat. Una forma d'actuar

¹³²En anglès s'utilitza el terme "proxy" que vol dir apoderat.

¹³³HUERGO LORA, A., "Una aproximación a los algoritmos...", *op. cit.*, 2020, pàg. 46 i 47.

¹³⁴BAROCAS, S. i SELBST, A. D., "Big Data's Disparate...", *op. cit.*, pàg. 691.

¹³⁵ZUIDERVEEN, F., *Discrimination, artificial intelligence...*, *op. cit.*, 2018, pàg. 21.

que, especialment en processos que afecten els drets de les persones, dificulta l'observança de transparència i publicitat dels criteris que han servit a la presa de decisions¹³⁶.

També és interessant el fenomen del lliscament, que opera en el sistema d'aprenentatge automàtic quan, al tenir prohibit emprar caràcters directament identificadors de categories protegides o indicadors propers, es desplaça l'atenció cap a indicadors més distants que en tot cas es refereixen a les categories malgrat ser més imprecisos (i a més, més difícilment identificables)¹³⁷.

2.1.5. Discriminació per assumpció (o adscripció) i per associació

La discriminació pot succeir a vegades per una assumpció sobre una persona (que pot ser o no ser correcte)¹³⁸. Pot passar habitualment en sistemes algorítmics quan s'elabora un perfil que assumeix (o adscriu) trets o característiques a persones o grups en base a, per exemple, dades de comportament. Les inferències que resulten de les dades de comportament poden ser o no ser correctes (a causa, entre altres, de manca de dades o errors), fent que s'assumeixin (o s'adscriuin) certes característiques sobre una persona que no siguin acurades (p. e. que té discapacitat) i es vegin discriminades per aquesta raó¹³⁹.

Altrament, la discriminació pot donar-se quan el sistema algorítmic selecciona les persones segons la seva associació amb persones que pertanyen a un grup desavantajat¹⁴⁰.

¹³⁶AÑÓN ROIG, M.J., “Desigualdades algorítmicas: conductas de alto riesgo para los derechos humanos”, *Derechos y libertades*, núm 47, època II, 2022, pàg. 30.

¹³⁷ZUDDAS, P., “Decisión algorítmica y discriminación”, *Revista General de Derecho Administrativo*, núm. 60, 2022, pàg. 6. “Lliscament” és la traducció que hem fet dels dos mots que apareixen en l'article original, “shift”, traduït per l'autor de l'article com “deslizamiento”.

¹³⁸ La Ley 15/2022, de 12 de julio, integral para la igualdad de trato y la no discriminación defineix en l'article 6.2.b la discriminació per error com aquella que es fonamenta en una apreciació incorrecte sobre les característiques de la persona o persones discriminades.

¹³⁹ GERARDS, J. i XENIDIS, R., *Algorithmic Discrimination in Europe...*, *op. cit.*, 2021, pàg. 68.

¹⁴⁰ WACHTER, S., “Affinity profiling and discrimination by association in online behavioural advertising”, *Berkeley Technology Law Journal*, vol. 35, núm. 2, 2020, pàg. 373. La Ley 15/2022, de 12 de julio, integral para la igualdad de trato y la no discriminación defineix en l'article 6.2.a la discriminació per associació com aquella que es dona quan una persona o grup en el qual s'integra, a causa de la seva relació amb una altra sobre la qual concorre alguna de les causes previstes en l'apartat primer de l'article 2 de la llei, és objecte de tracte discriminatori.

Aquesta manera de discriminar es pot veure incrementada a causa de les decisions preses mitjançant algorismes, atès que la seva tasca principal és realitzar associacions entre dades, no només aquelles que la persona decideix revelar, sinó també moltes d'altres que s'han pogut aconseguir a la xarxa. Les correlacions que puguin ser desconegudes per les persones es poden captar més fàcilment per un sistema automatitzat (per inferències en un conjunt de dades) i, per consegüent, poden amplificar la capacitat associativa amb finalitats pejoratives¹⁴¹.

Dins de la categoria de discriminació per associació es podria incloure el cas del sistema emprat per Facebook per excloure de l'accés a determinada publicitat a persones relacionades amb certs grups ètnics. Angwin i Parris Jr. van realitzar un estudi centrat en la publicitat sobre immobles que analitzava com el sistema de Facebook, mitjançant la inclusió d'una opció denominada "afinitats ètniques" en la configuració publicitària, donava la possibilitat d'excloure persones afroamericanes, asiàtiques-americanes i hispanes, és a dir, totes les comunitats no blanques¹⁴² de l'accés al contingut.

Així mateix, recuperant l'exemple dels codis postals i l'ètnia, a part de la discriminació arran de dades aparentment neutres, podem trobar un cas de discriminació per associació. Considerem que el 80% de les persones que viuen en el codi postal 08033 paga tard les seves factures. Si, basada en el perfil grupal, la companyia denega préstecs a totes les persones associades amb aquest codi postal, també denega els préstec al 20% restant que paga les factures a temps. Les prediccions serien incorrectes per les persones que tot i pagar les factures a temps es veurien incloses en el grup de persones no pagadores¹⁴³, essent discriminades per associació amb aquestes últimes.

En definitiva, cal remarcar que l'algoritme pot fer prediccions correctes en l'àmbit estadístic però que aplicades individualment sempre podran ser incorrectes. Imaginem que una companyia d'assegurances empra un model predictiu que agrupa a persones que condueixen un cotxe de marca "x", viuen en un codi postal "y" i són nascudes entre 1990 i 1996 en el "Grup 1", i determina que en aquest grup hi ha un 5% de possibilitats que les

¹⁴¹ RIVAS VALLEJO, P., "Sesgos de género en el uso de inteligencia artificial para la gestión de las relaciones laborales: análisis desde el derecho antidiscriminatorio", *e-Revista Internacional de la Protección Social (e-RIPS)*, vol. VII, núm. 1, 2022, pàg. 63.

¹⁴²ANGWIN, J. i PARRIS Jr., T., "Facebook lets advertisers exclude users by race", Propublica, 28 d'octubre 2016. Disponible a: <https://www.propublica.org/article/facebook-lets-advertisers-exclude-users-by-race>. Darrera visita el 2 de setembre de 2023.

¹⁴³ZUIDERVEEN, F., *Discrimination, artificial intelligence...*, op. cit., 2018, pàg. 68 i 69.

persones integrants del grup deixin a deure de mitjana 1000 euros a la companyia. Imaginem, llavors, que la companyia d'assegurances decideix incrementar el preu anual de l'assegurança per recuperar els diners que hipotèticament es poden perdre a causa dels impagaments. La predicció del model no diu la persona "z" del "Grup 1" no pagarà, sinó que estableix valors en base a tants per cent. Una persona a la qual l'algoritme integra en aquest grup haurà de pagar una prima més alta només perquè forma part del "Grup 1", essent això una diferenciació algorítmica controvertida¹⁴⁴.

2.1.6. Perpetuar estereotips en els quals es basen les estructures socials de discriminació

La discriminació algorítmica pot consistir també en retenir oportunitats o recursos¹⁴⁵, per exemple, denegació d'ajudes socials o préstecs o, en base a estructures socials de subordinació, perpetuar estereotips i dur a terme denigració cultural¹⁴⁶. No són pocs els casos en què s'ha detectat que els resultats obtinguts com a conseqüència de la introducció de determinades combinacions de paraules en els motors de recerca en la xarxa reproduïen els rols de gènere, cosifiquen les dones o contribueixen a la consolidació d'estereotips negatius sobre les persones que no són blanques¹⁴⁷. També l'edat cada vegada està més documentada com a factor de discriminació, per exemple, possibilitant en les plataformes de recerca de feina mitjançant algoritmes, l'exclusió de la visualització d'anuncis de campanyes de recerca de treballadors/es, a persones sol·licitants de feina majors d'una determinada franja d'edat¹⁴⁸.

També, per exemple, s'ha alertat dels efectes discriminatoris del sistema algorítmic basat en dades estadístiques del mercat laboral de l'Agència Estatal de Treball d'Àustria (AMS). Aquest sistema es dedica a classificar les persones que cerquen feina (analitzant diferents característiques personals i proporcionant una puntuació a cada una) en tres grups (segons

¹⁴⁴GERARDS, J. i ZUIDERVEEN, F., "Protected Grounds and the System...", *op. cit.*, 2022, pàg. 15.

¹⁴⁵ En anglès s'usen els termes "allocative harm" o "distributive harm", conceptes que es podrien traduir com perjudici d'assignació.

¹⁴⁶ En anglès s'usa el terme "representational harm", concepte que es podria traduir com perjudici de representació.

¹⁴⁷NOBLE, S. U., *Algorithms of Oppression: How Search Engines Reinforce Racism*, New York, New York University Press, 2018, pàg. 15, 16, 82 i 83.

¹⁴⁸AJUNWA, I., "Discrimination by platforms", *Berkeley Journal of Employment and labour law*, vol. 40, núm. 1, 2019, pàg. 9 i 10.

la probabilitat de trobar feina -alta, mitja o baixa-) amb l'objectiu d'assignar els recursos públics de la manera més eficient possible¹⁴⁹. En els dos primers grups -integrats per persones que suposadament tindran més opcions d'entrar al mercat laboral-, es destinen recursos com assessorament individual i accés a programes de formació professional, però en canvi, a persones del tercer grup -les opcions de les quals en el mercat laboral l'algoritme defineix com a baixes-, se'ls ofereix un suport diferent d'acord amb el seu perfil. Això pot ocasionar que la privació de recursos condueixi a la realització i validació de la predicció que són persones menys proclius a trobar feina i per tant no s'ha d'invertir recursos en elles¹⁵⁰. S'ha acusat el sistema d'assignar persones que formen part de grups prèviament discriminats i marginats (p. e. dones) a un grup classificatori menys favorable (p. e. persones que cerquen feina des de fa molt temps), on rebran menys suport. El govern Austríac ha reconegut que l'algoritme ha mostrat que les responsabilitats de cura afecten les oportunitats en el mercat de treball de les dones però no dels homes. De manera similar, l'eina "My employability" a Croàcia emprà l'edat dels fills de les persones usuàries com a rellevant en relació amb l'ocupabilitat. Quan l'edat és de 0-2 anys s'ha estipulat que té un impacte negatiu en les prediccions de trobar feina de les dones, mentre que pels homes el paràmetre no es té en compte. Tot i ser només una eina de càlcul de probabilitats, desencoratja dones amb fills petits a entrar en el mercat laboral¹⁵¹.

Gerards i Zuiderveen alerten també del risc que la diferenciació algorítmica genera al reforçar les desigualtats estructurals entre grups socioeconòmics. Les botigues en línia poden cobrar diferents preus pels mateixos productes a persones consumidores diferents en base a la informació que la botiga té sobre la persona consumidora: diferenciació de preu en línia. Per exemple, una botiga als Estats Units d'Amèrica del Nord cobrava preus més alts a persones que vivien en pobles que a persones que vivien en les ciutats (a Estats Units d'Amèrica del Nord les persones que viuen en grans ciutats són, de mitjana, més riques que les persones que viuen al camp), atès que en les ciutats hi havia més competència d'altres botigues amb els mateixos productes (cobrant-los preus més baixos

¹⁴⁹OECD, *Profiling tools for early identification of jobseekers who need extra suport*, Paris, 2018, pàg. 3.

¹⁵⁰ALLHUTTER, D., CECH, F., FISCHER, F., GRILL, G. i MAGER, A., "Algorithmic Profiling of Job Seekers in Austria: How Austerity Politics Are Made Effective", *Frontiers Big Data*, vol. 3, art. 5, 2020, pàg. 7.

¹⁵¹GERARDS, J. i XENIDIS, R., *Algorithmic Discrimination in Europe...*, op. cit., 2021, pàg. 86.

per evitar que s'adrecessin a la competència) i en canvi, en les àrees rurals, en estar les botigues de la competència a hores de distància, podien cobrar més diners¹⁵².

Dos exemples del procés que segueix l'algoritme per arribar a resultats discriminatoris que reforcen desigualtats estructurals són els següents. Talent Match és una funció de LinkedIn que recomana possibles persones treballadores a persones que contracten. Si LinkedIn basa les seves recomanacions en l'interès demostrat per part d'aquelles persones que contracten per cert tipus de persones candidates, les recomanacions que oferirà reflectiran qualsevol biaix que tinguin. Si els algoritmes observen que desvaloren certes persones candidates que formen part d'una categoria sensible, disminuiran el ritme amb el qual les recomanen, aprenent a reproduir les seves preferències perjudicials¹⁵³. Les prediccions, d'aquesta manera, poden arribar a autoafirmar-se, esdevenint un bucle que es perpetua, com quan les dades històriques contenen discriminacions ja sigui en les pràctiques que les creen o en la recollida de dades, existint el risc que un procés objectiu esdevingui un mitjà per perpetuar discriminació històrica¹⁵⁴.

També ens podem trobar aquest tipus de discriminació en el cas de les opinions donades per part de les persones consumidores sobre les persones que els han prestat un servei. Actualment és crucial per a les empreses el seu prestigi i posicionament en xarxes socials, per això s'apliquen mecanismes reputacionals per obtenir la resposta de la clientela després de la prestació del servei, de la qual depèn la reputació empresarial digital i la competitivitat de l'empresa. Les persones consumidores es converteixen en còmplices de la construcció de la marca d'empresa ja que aquests mecanismes faciliten l'entrada i publicitat d'opinions de la clientela, implicant això que qualsevol crítica, positiva o negativa, pugui tenir efecte multiplicador sobre la marca empresarial en xarxes socials.

¹⁵² GERARDS, J. i ZUIDERVEEN, F., "Protected Grounds and the System...", *op. cit.*, 2022, pàg. 12 i 13.

¹⁵³ WOODS, D., "LinkedIn's Monica Rogati on "What Is a Data Scientist?""", *Forbes*, 27 de novembre de 2011. Disponible a: <https://www.forbes.com/sites/danwoods/2011/11/27/linkedin-monica-rogati-on-what-is-a-data-scientist/>. Darrera visita el 2 de setembre de 2023.

¹⁵⁴ BENNET MOSES, L. i CHAN, J., "Algorithmic prediction in policing...", *op. cit.*, 2018, pàg. 810. Cal fer esment obligat a l'anomenada policia predictiva, un sistema de predicció de les zones on es probable que es cometin més crims. Aquesta predicció ocasiona que s'incrementi la vigilància en aquests espais on augmentarà, consegüentment, la detecció de crims, creant un bucle que s'autoafirma. Per exemple, als Països Baixos, un sistema de policia predictiva "Crime Anticipation System" (CAS), per predir el risc de la comissió de crims en base a l'anàlisi algorítmic de dades d'atrestats policials. En relació amb la policia predictiva, en la doctrina espanyola, vegeu: DÍAZ GONZÁLEZ, G.M., "Algoritmos y actuación policial: la policia predictiva", a HUERGO LORA, A.J., (dir.) i DÍAZ GONZÁLEZ, G. M. (coord.), *La regulación de los algoritmos*, *op. cit.*, 2020, pàg. 181-234.

En alguns casos això pot condicionar la configuració del servei prestat en relació amb les persones que el presten, com és el cas de les plataformes d'“economia col·laborativa”¹⁵⁵.

El Tribunal de Bolonya, en sentència de 31 de desembre de 2020¹⁵⁶, qualifica com a discriminador l'algoritme “Frank” emprat per l'empresa Deliveroo per a la selecció de persones treballadores donat l'ús de mecanismes de reputació motivats en les opinions de la clientela i en l'absència de consideració de factors justificatius de la cancel·lació de serveis prèviament reservats per les pròpies persones treballadores¹⁵⁷. L'algoritme, quan elabora els rànquings de reputació de les persones ciclistes, realment determina les oportunitats de treball futur i les reserves prioritàries de repartiment, marginant fins el punt d'excloure aquelles persones que no poden estar sempre disponibles en els horaris que se'ls atorga o que no s'adapten a la lògica de l'algoritme, és a dir, que aquelles persones treballadores que la seva disponibilitat no els permet acceptar totes les feines que els proposa l'algoritme es veuen desavantatjades en les subsegüents feines i hi ha risc que se'ls exclogui completament de les oportunitats de feina¹⁵⁸. Les dones són un col·lectiu víctima d'aquesta discriminació, ja que els desequilibris estructurals en els rols de gènere que fan que les dones de mitjana dediquin més temps a tasques de cura que els homes, ocasionen que estiguin menys disponibles per treballar hores flexibles a demanda, perquè han de combinar treball pagat amb tasques de cura¹⁵⁹, veient-se a causa d'aquest fet, discriminades per l'algoritme.

Finalment, no s'ha d'obviar tampoc, com ja criticava Adam l'any 1998, que el desenvolupament per part d'homes de la intel·ligència artificial, implica passar per alt les dones com a dissenyadores d'aquesta tecnologia, inscrivint-se en aquesta una visió

¹⁵⁵RIVAS VALLEJO, P., “Discriminación algorítmica: detección, prevención y tutela”, *op. cit.*, 2021, pàg. 25 i 26. L'autora posa l'exemple següent: part de la clientela d'una empresa que presta serveis de transport amb cotxe, té conviccions estereotipades sobre les dones, i els crea malestar que una dona condueixi el cotxe en el servei contractat de manera que transmet el seu malestar a l'empresa que el presta (mitjançant opinions i observacions). És possible que l'empresa condicionei la seva política de contractació en base a les preferències estereotipades de la clientela.

¹⁵⁶ Tribunal de Bolonya, sentència núm. 2949/2020, de 31 de desembre de 2020. Disponible a: <https://www.algoritmolegal.com/wp-content/uploads/2021/01/Sentencia-Bologna-Italia-Deliveroo-dic-2020-Original-italiano.pdf> Darrera visita el 2 de setembre de 2023.

¹⁵⁷ RIVAS VALLEJO, P., “Discriminación algorítmica: detección, prevención y tutela”, *op. cit.*, 2021, pàg. 37.

¹⁵⁸ En el sentit del control per part de l'empresa de les oportunitats de feina de les persones treballadores en el cas de la plataforma Glovo, vegeu: Tribunal Supremo (Sala de lo Social, Pleno), sentència 805/2020, de 25 de setembre de 2020 (recurso 4746/2019), FJ 21.1 i 21.2.

¹⁵⁹ GERARDS, J. i XENIDIS, R., *Algorithmic Discrimination in Europe...*, *op. cit.*, 2021, pàg. 56, 83.

esbiaixada del món¹⁶⁰. Torralba planteja de manera molt acurada que en els sistemes d'intel·ligència artificial, el passat es converteix completament en el garant del futur, tenint l'efecte pernicios de restringir l'obertura de possibilitats per a aquells que ja estan confinats als marges, a més d'oferir més oportunitats als qui ja són privilegiats¹⁶¹.

2.2. Discriminació en base a apreciacions humanes

Podem classificar en aquest apartat la discriminació causada deliberadament per persones que empren el sistema algorítmic per emmascarar la seva responsabilitat (2.2.1) i aquella arran de les decisions que es prenen en base als resultats dels sistemes algorítmics (2.2.2).

2.2.1. Ús dels sistemes algorítmics per emmascarar la discriminació

L'anàlisi massiva de dades pot incorporar noves maneres de discriminar intencionadament ja que aquelles persones que vulguin fer-ho, només han d'explotar totes les mesures tècniques que hem descrit per obtenir resultats esbiaixats. És a dir, que tota forma de discriminació que es dona de manera no intencional, també es pot donar de manera intencional¹⁶².

Les persones que prenen les decisions podrien de manera deliberada esbiaixar la recollida de dades per assegurar normes menys favorables per a membres de grups vulnerables, o podrien amagar els efectes perjudicials d'anteriors sistemes de presa de decisions insistint en què constitueixen un conjunt d'exemples fiables i imparcials¹⁶³.

¹⁶⁰ ADAM, A., *Artificial knowing: Gender and the thinking Machine*, Florence, KY, USA, Routledge, 1998, pàg. 1.

¹⁶¹ TORRALBA, F., *L'ètica algorítmica*, op. cit., 2022, pàg. 78. Vegeu també: PONCE SOLÉ, J., "Las relaciones entre inteligencia artificial, regulación y ética, con especial atención al sector público", *Revista General de Derecho Administrativo*, núm. 61, 2022, pàg. 2 i 3.

¹⁶² BAROCAS, S. i SELBST, A. D., "Big Data's Disparate...", op. cit., 2016, pàg. 692.

¹⁶³ *Ídem, ibidem*.

Com observen Kroll *et al.* existeix el problema d'“emascarar”, és a dir, quan la discriminació intencional es disfressa a través d'una forma de discriminació no intencional, per exemple, una persona que pren decisions amb prejudicis podria esbiaixar les dades d'entrenament o escollir indicadors per a categories protegides amb la intenció de generar resultats discriminatoris¹⁶⁴.

2.2.2. Decisions que es prenen en base als resultats dels sistemes algorítmics

Existeix també la possibilitat de que no sigui el propi algoritme el que discrimini sinó que les decisions que es prenguin en base als resultats que genera siguin discriminatòries. Per exemple, en relació amb l'AMS, ja esmentat, l'empresa contractada va publicar informació sobre el funcionament de l'algoritme i s'evidenciava que es restaven punts a aquelles persones que pertanyien a grups desavantatjats o vulnerables, i per consegüent, més susceptibles de patir discriminació. Malgrat aquest sistema es limitava a reflectir les pràctiques discriminatòries existents en el mercat laboral, contribuïa a estigmatitzar aquests grups i els seus membres classificant-los, pel mer fet de ser més vulnerables, en una categoria inferior. No obstant això, el servei públic austríac va defensar la necessitat d'aquest sistema per proporcionar una ajuda més adequada a les persones que podien trobar més dificultats per accedir al mercat laboral¹⁶⁵. Es va arribar a la conclusió (és a dir, es va prendre la decisió en base al resultat de l'algoritme), que l'assignació de recursos més eficaç seria donar més recursos a persones amb probabilitat “mitja” de trobar una nova feina i donar-ne menys a aquelles amb probabilitat “baixa”, perpetuant així les situacions d'exclusió social que pateixen certs grups de persones, i reforçant la construcció d'estructures i institucions socials mitjançant relats de subordinació dels grups històricament oprimits¹⁶⁶.

També, en aquest mateix sentit, s'està augmentant l'ús de sistemes algorítmics privats per detectar el frau en la sol·licitud d'ajudes públiques¹⁶⁷, procés que està allunyant els

¹⁶⁴KROLL, J.A. *et al.*, “Accountable algorithms...”, *op. cit.*, 2017, pàg. 682.

¹⁶⁵SORIANO ARNANZ, A., “Decisiones automatizadas y discriminación...”, *op. cit.*, 2021, pàg. 13.

¹⁶⁶*Ídem*, pàg. 13-14.

¹⁶⁷ALSTON, P., *Digital welfare states and human rights*, UN Special Rapporteur on extreme poverty and human rights, report A/74/493, New York, 2019, pàg. 10-11.

sistemes de seguretat social i els programes d'ajudes públiques del seu objectiu d'alleujar la pobresa i donar suport a la ciutadania¹⁶⁸. S'entén que el que es busca amb l'ús d'algoritmes és una distribució més eficient dels recursos, no obstant això, moltes vegades les decisions preses només en base a factors quantitius evidencien la manca de valor que es dona als factors qualitius, que són aquells que efectivament tenen incidència en la vida de les persones.

Tot seguit analitzem, arran de les formes de discriminació anteriorment descrites, quins factors caracteritzen la discriminació ocasionada mitjançant sistemes algorítmics, i si l'existència d'aquests factors n'agreuja l'efecte.

3. Factors que caracteritzen la discriminació algorítmica

Després d'enumerar les diferents formes de discriminar que podem trobar en les decisions automatitzades mitjançant sistemes algorítmics, és pertinent abordar els riscos i problemes que s'afegeixen, a causa de la complexitat d'aquesta eina, en la lluita contra la discriminació, a saber, en primer lloc, els biaixos que podem trobar en els algoritmes (3.1), en segon lloc, l'opacitat intrínseca a aquests sistemes (3.2), en tercer lloc, el grau de personalització que poden aconseguir (3.3) i, en quart lloc, la seva capacitat per magnificar els efectes discriminatoris (3.4).

3.1. Els biaixos algorítmics com a font de la discriminació

Els biaixos en la presa de decisions tenen una incidència directa en la discriminació. Principalment esdevé controvertit que els biaixos poden ser causats de moltes maneres¹⁶⁹

¹⁶⁸RANCHORDÁS, S. i SCHUURMANS, Y., "Outsourcing the welfare state: the role of private actors in welfare fraud investigations", *European Journal of Comparative Law and Governance*, vol. 7, núm. 1, 2020, pàg. 10.

¹⁶⁹SPRINGER, A., GARCIA-GATHRIGHT, J. i CRAMER, H., "Assessing and Addressing Algorithmic Bias— But Before We Get There", Association for the Advancement of Artificial Intelligence, 2017, pàg. 2. Es planteja en l'article l'existència de tres tipologies de biaixos algorítmics: biaixos algorítmics en les dades

i que fan que la decisió que pren el sistema automatitzat, de manera molt subtil i subreptícia, pugui ser discriminatòria. Els biaixos ocults en ocasions obeeixen a un objectiu clarament discriminatori, però, en la majoria dels casos, simplement són provocats pel simple desinterès cap al seu impacte col·lateral¹⁷⁰.

Podem trobar en la literatura existent nombrosos exemples sobre l'existència de diferents biaixos algorítmics relacionats amb el procés de contractació¹⁷¹, el contingut i l'ús de pàgines web¹⁷² -concretament en relació amb col·lectius de població històricament racialitzats desfavorits¹⁷³-, els sistemes de reconeixement facial que tendeixen a no identificar de forma igualment precisa a les persones negres¹⁷⁴, o els sistemes d'IA que pretenen identificar l'orientació sexual d'una persona¹⁷⁵.

Es defineix biaix com una inclinació que afavoreix o perjudica a una persona, objecte o posició¹⁷⁶. És a dir, en el cas de biaixos discriminatoris, una desviació inadequada en el procés d'inferència¹⁷⁷. Són riscos inherents a qualsevol activitat econòmica i social i es poden donar tant en decisions preses per éssers humans com en decisions preses per sistemes algorítmics. No obstant això, el mateix biaix en un sistema algorítmic pot tenir un efecte molt més ampli (problema que abordarem específicament en l'apartat 3.4 d'aquest capítol), afectant i discriminant moltes persones sense els mecanismes de control social que governen el comportament humà¹⁷⁸. Una persona pot de manera deliberada o

introduïdes, biaixos computacionals que resultin de l'algoritme o de les decisions de l'equip que gestiona el funcionament del sistema, o biaixos de resultat per a grups de persones usuàries específiques.

¹⁷⁰ RIVAS VALLEJO, P., "Análisis desde el derecho antidiscriminatorio", a RIVAS VALLEJO, P., (dir.), *Discriminación algorítmica en el ámbito laboral: perspectiva de género e intervención*, Navarra, Aranzadi, 2022, pàg. 437; RIVAS VALLEJO, P., "Sesgos de género en el uso...", *op. cit.*, 2022, pàg. 54.

¹⁷¹Vegeu en aquest sentit: CHEN, L., MA, R., HANNÁK, A. i WILSON, C., "Investigating the Impact of Gender on Rank in Resume Search Engines", *Annual Conference of the ACM Special Interest Group on Computer Human Interaction*, 2018, pàg. 1-14.

¹⁷²Vegeu en aquest sentit: BAEZA-YATES, R., "Bias on the Web", *Communications of the ACM*, vol. 61, núm. 6, 2018, pàg. 54-61.

¹⁷³Vegeu en aquest sentit: SWEENEY, L., "Discrimination in Online Ad Delivery", *Acm Queue*, vol. 11, núm. 3, 2013, pàg. 10-29.

¹⁷⁴Vegeu en aquest sentit: BUOLAMWINI, J. i GEBRU, T., "Gender Shades: Intersectional Accuracy Disparities in Commercial Gender Classification", *Proceedings of the 1st Conference on Fairness, Accountability, and Transparency*, vol. 81, 2018, pàg. 77-91.

¹⁷⁵Vegeu en aquest sentit: WANG, Y. i KOSINSKI, M., "DeepNeural Networks Are More Accurate Than Humans at Detecting Sexual Orientation from Facial Images" *Journal of Personality and Social Psychology*, vol. 114, núm. 2, 2018, pàg. 246-257.

¹⁷⁶GRUPO DE EXPERTOS DE ALTO NIVEL SOBRE INTELIGENCIA ARTIFICIAL, *Directrices éticas para una IA fiable*, Comisión Europea, Bruselas, 8 de abril de 2019, considerando 149.

¹⁷⁷AEPD, *Adecuación al RGPD de tratamientos que incorporan Inteligencia Artificial. Una introducción*, Madrid, 2020, pàg. 7.

¹⁷⁸EUROPEAN COMMISSION, *White paper on artificial Intelligence. A European Approach to excellence and trust*, COM(2020) 65 final, Brussels, 19 February 2020, pàg. 11-12.

de manera involuntària discriminar, però una màquina no té intencionalitat. La confiança que es deposita en ella radica en la consideració que els càlculs algorítmics que duu a terme són objectius i els resultats que produeix estan basats en les dades disponibles¹⁷⁹.

Els biaixos poden sorgir en els sistemes d'intel·ligència artificial de nombroses maneres. En els sistemes impulsats per dades o creats a través de l'aprenentatge automàtic, els biaixos poden ocasionar-se en la recollida de dades i en la formació del sistema. També poden sorgir en els sistemes lògics o basats en normes com a conseqüència de la visió que pot tenir un/a enginyer/a de les regles aplicables en un entorn específic. O, a causa de la formació i adaptació en línia a través de la interacció, o com a conseqüència de la personalització en aquells casos en què es presenten als usuaris recomanacions o informacions adaptades als seus gustos¹⁸⁰. No han d'estar relacionats necessàriament amb inclinacions humanes o amb la recollida de dades per part de les persones. Poden sorgir, per exemple, en els limitats contextos en els que s'utilitza un sistema, no existint així la possibilitat de generalitzar-lo a un altre context. Poden ser positius o negatius, intencionats o no intencionats, donant lloc, en alguns casos, a resultats discriminatoris o injustos¹⁸¹, i són particularment greus quan, per exemple, deriven de discriminacions d'un grup en favor d'un altre¹⁸². També poden sorgir, i és rellevant diferenciar-los, o bé a causa d'un entrenament en base a decisions preses per persones, o bé a causa d'un entrenament en base a les disparitats que es donen en la població, és a dir, aprenent a reproduir resultats que són producte de l'opressió històrica¹⁸³. En els casos on el resultat no s'ha pogut prevenir o anticipar en la fase de disseny, els riscos es derivaran dels impactes pràctics de les correlacions o patrons que el sistema identifica en el conjunt de dades que té a l'abast¹⁸⁴. En la interpretació dels resultats de sistemes algorítmics, també es poden donar els biaixos quan s'accepten els resultats com a certs i inamovibles, sense esperit crític, assumint un principi d'autoritat derivat de les expectatives creades pel sistema¹⁸⁵.

¹⁷⁹LEESE, M., "The New Profiling: Algorithms, Black Boxes, and the Failure of AntiDiscriminatory Safeguards in the European Union", *Security Dialogue*, vol. 45, núm. 5, 2014, pàg. 502-503.

¹⁸⁰GRUPO DE EXPERTOS DE ALTO NIVEL SPBRE LA INTELIGENCIA ARTIFICIAL, "Directrices éticas...", *op. cit.*, 2019, pàg. 48-49.

¹⁸¹*Ídem, ibidem.*

¹⁸²AEPD, *Adecuación al RGPD de tratamientos que incorporan Inteligencia Artificial. Una introducción*, Madrid, 2020, pàg. 7-8.

¹⁸³GREEN, B, "The False Promise...", *op. cit.*, 2020, pàg. 600.

¹⁸⁴EUROPEAN COMMISSION, *White paper on artificial Intelligence. A European Approach...*, *op. cit.*, 19 February 2020, pàg. 11-12.

¹⁸⁵AEPD, *Adecuación al RGPD de tratamientos que incorporan Inteligencia Artificial. Una introducción*, Madrid, 2020, pàg. 7-8.

En un sistema algorítmic, llavors, un biaix és un error sistemàtic de qualsevol tipus en el resultat de les operacions algorítmiques i té un significat molt més ampli que discriminació, ja que no només té a veure amb errors injustos sinó que té a veure amb tots els tipus d'errors sistemàtics, que poden incloure aquells de natura estadística, cognitiva, estructural o institucional¹⁸⁶. Quan s'evoca en el context de l'equitat, el biaix algorítmic es refereix a una tipologia d'error particular que posiciona grups privilegiats en avantatge sistemàtic i grups no privilegiats en desavantatge sistemàtic¹⁸⁷. El terme biaix algorítmic engloba més que el terme discriminació algorítmica ja que és refereix a qualsevol tipus de desavantatge que es podria veure com èticament o moralment censurable. Per exemple, un algoritme que perjudica grups amb poc poder adquisitiu i beneficia persones amb alt poder adquisitiu es podria veure com una forma de discriminació algorítmica des d'una perspectiva ètica¹⁸⁸ (també pot passar això en base a algoritmes que sistemàticament perjudiquen persones que es troben en una situació d'atur de llarga duració, una de les característiques criticades a l'AMS); però des d'una perspectiva legal, per exemple, la discriminació algorítmica només respon al tractament injustificat i desfavorable o al desavantatge viscut per categories específiques de la població protegides per la llei¹⁸⁹.

Els dos elements més rellevants quant als biaixos, i que abordarem seguidament de manera separada, són les dades personals amb les quals s'arriba a les conclusions esbiaixades (3.1.1) i la intencionalitat de la persona que pot crear els biaixos (3.1.2).

3.1.1. Irrellevància de la categorització de dades “sensibles” quan ens referim a les decisions automatitzades

¹⁸⁶ Sobre la diferència entre biaix i discriminació, vegeu: EUROPEAN UNION AGENCY FOR FUNDAMENTAL RIGHTS (FRA), *Bias in Algorithms - Artificial Intelligence and Discrimination*, Luxembourg, 8 December 2022, pàg. 22-25.

¹⁸⁷ BELLAMY, R. K. E. *et al.*, “AI Fairness 360: An Extensible Toolkit for Detecting, Understanding, and Mitigating Unwanted Algorithmic Bias”, 2018, pàg. 2. Disponible a: <https://arxiv.org/>. Darrera visita el 2 de setembre de 2023.

¹⁸⁸ Veiem aquí com la discriminació causada per sistemes algorítmics està estretament relacionada amb el debat sobre l'ètica en la intel·ligència artificial. Vegeu, en aquest sentit: LOBACHEVA, A. i KASHTANOVA, E., “Social Discrimination in the epoch...”, *op. cit.*, 2022, pàg. 97-105.

¹⁸⁹ GERARDS, J. i XENIDIS, R., *Algorithmic Discrimination in Europe...*, *op. cit.*, 2021, pàg. 47.

L'objectiu d'aquest apartat és clarificar com l'anàlisi massiva de dades suposa un nou risc en relació a la discriminació degut a que la quantitat de dades i la qualitat del seu anàlisi fa que totes elles puguin esdevenir sensibles¹⁹⁰.

La “dataficació”, neologisme per referir-nos a la recopilació, anàlisi i intercanvi sistemàtic de dades, així com els resultats obtinguts, suposa que el tractament de dades i les tècniques avançades d'anàlisi són capaces de convertir literalment qualsevol cosa en dades, incloses aquelles que mai hem pensat que podrien contenir-ne. Se'n desprèn com a conseqüència, que és possible fer connexions entre persones i dades més ràpid i de manera més variada a com mai s'havia fet. Això ocasiona tant que hi hagi un increment de la quantitat de dades disponibles, com que més dades siguin personals, essent el canvi tant qualitatiu com quantitatiu¹⁹¹.

Des del Convenio núm. 108 del Consejo de Europa para la protección de las personas con respecto al tratamiento automatizado de datos de carácter personal¹⁹² s'han regulat les dades especialment protegides, les dades sensibles o les categories especials de dades, normalment prohibint el seu tractament llevat que s'asseguessin certes condicions de llibertat en donar consentiment o sobre la base de prevalença de valors superiors com la salut o la preservació de la vida. Com explica Martínez, aquestes categories de dades s'han objectivat i responen a una perspectiva estàtica, deixant desatesa la perspectiva des d'un punt de vista qualitatiu (explica que aquest últim punt de vista exigeix considerar la sensibilitat des de la possible repercussió material d'aquestes dades en l'esfera dels drets de les persones). Apunta que el processament massiu de dades aporta unes possibilitats de modelatge de la informació capaces de convertir en sensibles dades que no responen objectivament a aquesta categorització segons la llei¹⁹³. Posa l'exemple dels processos de selecció de personal, en els quals es podrien identificar patrons de conducta mitjançant l'anàlisi de quantitats ingents de dades generades per la pròpia persona en entorns socials.

¹⁹⁰ SCHREURS, W., “Cogitas, Ergo Sum. The role of data protection Law and non-discrimination law in group profiling in the private sector”, a HIDELEBRANDT, M. i GUTWIRTH, S., (eds.), *Profiling the European citizen: Cross-Disciplinary Perspectives*, Berlin, Springer Dordrecht, 2008, pàg. 245; CRAWFORD, K. i SCHULTZ, J., “Big data and due process...”, *op. cit.*, 2014, pàg. 102. S'estableix en aquest article que l'anàlisi predictiva pot esdevenir informació personal identificable i aquest procés pot predir informació altament íntima, malgrat cap de les dades inicials es pugui considerar com personal.

¹⁹¹PURTOVA, N., “Do property rights in personal data make sense after the big data turn? Individual control and transparency”, *Journal of Law and Economic Regulation*, vol. 10, núm. 2, 2017, pàg. 66.

¹⁹²Convenio núm. 108 del Consejo de Europa para la protección de las personas con respecto al tratamiento automatizado de datos de carácter personal, hecho en Estrasburgo el 28 de enero de 1981.

¹⁹³MARTÍNEZ, R., “Cuestiones de ética jurídica al abordar proyectos de Big Data. El contexto del Reglamento General de Protección de Datos”, *Dilemata*, núm. 24, 2017, pàg. 157-158.

Aquestes dades podrien fer arribar a la conclusió que la persona desenvolupa conductes que poden determinar riscos per a la salut, creant dades qualificades com d'especial protecció però en canvi, no protegides per la legislació¹⁹⁴.

Per consegüent, la remissió d'informació a categories de dades concretes resulta problemàtica. L'especificitat és merament superficial ja que és altament probable que les categories més concretes no refereixin de manera detallada tots els subtipus d'informació que es poden situar dins d'elles. En conseqüència, tots els conceptes que es poden integrar en aquestes categories però no estan explícitament regulats, ocasionen una ambivalència que es tradueix en buits legals, per exemple en relació amb les dades que revelen la condició de ser fumador¹⁹⁵ o les històries clíniques de persones que han mort¹⁹⁶, entre moltes d'altres¹⁹⁷.

Donada la crítica que se li ha fet a la categorització de dades sensibles, s'ha plantejat també si eliminant l'existència d'aquestes dades en l'anàlisi massiva s'arribaria a resultats menys discriminatoris o, al contrari, es provocarien resultats encara més esbiaixats¹⁹⁸. Es detalla en l'estudi de Williams, Brooks i Shmargad que quan s'enregistren com a "dades" els rastres de les vides de les persones, el resultat està ple de correlacions, és a dir, característiques personals que tendeixen a aparèixer juntes, a causa de la repetició d'hàbits i de la presa de decisions en els mateixos sistemes socials i influències compartides amb altres persones¹⁹⁹. Per tant, si se suprimeixen dades referents, per exemple, a l'ètnia o al

¹⁹⁴*Ídem*, pàg. 158.

¹⁹⁵Qüestió objecte d'un informe jurídic (0129/2005) de l'AEPD anomenat "*Naturaleza del dato de fumador como dato de salud*".

¹⁹⁶Qüestió objecte d'un informe jurídic (0149/2008) de l'AEPD anomenat "*Datos de salud de personas fallecidas*".

¹⁹⁷HUERTA ANGUIANO, J.A., "“Naturaleza Intrínseca”, “Contexto” o “Finalidad” en la determinación del carácter sensible de los datos personales", *Estudios en derecho a la información*, núm. 10, 2020, pàg. 11.

¹⁹⁸Actualment, molts/es autors/es es decanten per la inclusió de categories protegides en el tractament de dades per millorar la detecció de la discriminació, vegeu, per exemple: HOFFMANN, H., VOGT, V., HAUER, M. P. i ZWEIG, K., "Fairness by awareness? On the inclusion of protected features in algorithmic decisions", *Computer, Law and Security Review*, vol. 44, 2022, pàg. 11; HOLSTEIN, K. *et al.*, "Improving Fairness in Machine Learning Systems: What Do Industry Practitioners Need?", *Proceedings of the ACM Conference on Human Factors in Computing Systems*, 2019, pàg. 1-16; ANDRUS, M. *et al.*, "What We Can't Measure, We Can't Understand: Challenges to Demographic Data Procurement in the Pursuit of Fairness", *Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency*, 2021, pàg. 249-260. En l'àmbit espanyol, vegeu: SORIANO ARNANZ, A., "Decisiones automatizadas y discriminación...", *op. cit.*, 2021, pàg. 38; SORIANO ARNANZ, A., "La propuesta de Reglamento de Inteligencia Artificial de la UE y los sistemas de alto riesgo", *Revista General de Derecho de los Sectores Regulados*, núm. 8, 2021, pàg. 16.

¹⁹⁹WILLIAMS, B. A., BROOKS, C. F. i SHAMARGAD, Y., "How Algorithms Discriminate Based on Data they Lack: Challenges, Solutions, and Policy Implications", *Journal of Information Policy*, vol. 8, 2018, pàg. 82. Un bon exemple que es dona sobre aquest fenomen es troba en el sector del treball, en l'anomenada

sexe, el resultat de les correlacions d'informació entès com a “dades d'una persona” serà menys precís i estarà més construït en base a patrons i correlacions trobades en persones amb característiques similars. Quan ens referim a macro dades, aquest fenomen s'accentua, més camps estan omplerts a través de sistemes de predicció de manera que mantenir aspectes centrals de la identitat en privat no sempre pot prevenir que aquesta informació sigui inferida, amb una base molt menys precisa, que dificultaria la identificació de discriminacions a través d'auditories, impossibilitant arribar a conclusions on l'existència de categories de dades “sensibles” és imprescindible²⁰⁰.

Seguidament analitzarem com l'anàlisi de dades, no pot ser neutra, ja que sigui intencionalment o no, sempre reproduïx biaixos existents en la societat.

3.1.2. Intencionalitat en la presa de decisions. L'anàlisi de dades mai és neutra

L'ús de sistemes algorítmics ha comportat preocupacions en relació als potencials biaixos no intencionals i no desitjats que poden crear, essent aquests inescrutables per a les persones que es veuen afectades. Per exemple, un algoritme de l'empresa “Google” que presumptivament dirigia a homes més que a dones ofertes de feina on el salari era més alt²⁰¹. Si bé és cert que podria ser que de manera intencional s'haguessin establert aquests criteris, és molt probable que s'establís aquesta dinàmica, ja que en les dades d'entrenament dels algoritmes i que aquests reproduïen, es reflectia la societat actual on els homes de manera generalitzada tenen sous més alts que les dones. És pertinent remarcar d'aquest exemple, que un risc afegit en la detecció de la discriminació també ve derivat dels nous anuncis digitals, que es caracteritzen per tenir la virtualitat d'interactuar única i exclusivament amb cada persona usuària de la plataforma digital on s'ubiquen, així com d'impedir que les persones excloses de la seva visualització en siguin

“discriminació estadística”, que ocorre quan les persones que contracten no tenen suficient informació sobre les habilitats d'una persona candidata i omplen els detalls que manquen en base als seus coneixements sobre anteriors persones candidates.

²⁰⁰*Ídem*, pàg. 83-90.

²⁰¹DATTA, A., TSCHANTZ, M. C. i DATTA, A., “Automated Experiments on Ad Privacy Settings: A Tale of Opacity, Choice and Discrimination”, *Proceedings on Privacy Enhancing Technologies*, 2015, pàg. 93 i 102.

conscients²⁰² -contràriament als anuncis tradicionals que no impedeixen la seva visualització per part de qualsevol persona.

Un altre exemple es troba en què, conferint més atenció i oportunitats a aquelles persones treballadores que es preveu que seran més competents en relació a alguna tasca, aquells que contracten es trobaran que persones més vulnerables socialment es veuen constantment sotmeses a tractaments menys avantatjosos perquè el criteri per contractar empleats vulnerables socialment és sistemàticament pitjor. Això pot passar fins i tot si els criteris prohibits s'han eliminat de les bases de dades, les dades són lliures de prejudicis i biaixos, les característiques són diverses i l'únic objectiu és maximitzar la precisió classificatòria o predictiva²⁰³. Per tant, biaixos arbitraris poden sorgir fins i tot quan els judicis no són explícitament perjudicials i poden autopropaguar-se quan qui pren la decisió actua en base a aquests judicis esbiaixats²⁰⁴.

S'atribueix aquest tipus de biaixos a dades mal seleccionades, incompletes, incorrectes o antiquades, perpetuació no intencional i promoció de biaixos històrics (introduïts en les dades), sistemes de correlació mal dissenyats, serveis de personalització i recomanació que redueixen en comptes d'expandir les opcions de les persones usuàries, sistemes de presa de decisió que assumeixen que correlació necessàriament suposa causa, o conjunts de dades que els falta informació o que representen de manera desproporcionada certs grups de població²⁰⁵. És evident llavors, que a causa dels biaixos ocults en el model predictiu, com dades no representatives o exemples mal categoritzats, el model arriba a resultats discriminatoris²⁰⁶. Aquí és on pren importància la facilitat amb la qual la discriminació pot ser oculta o ocultada en la presa de decisions mitjançant algorismes, ja que hi ha molts factors que poden influir en què un resultat sigui injustament discriminatori²⁰⁷ i tant poden ser intencionals com no intencionals. Aquest fet és

²⁰²RIVAS VALLEJO, P., “Discriminación algorítmica: detección, prevención y tutela”, *op. cit.*, 2021, pàg. 30 i 31.

²⁰³BAROCAS, S. i SELBST, A. D., “Big Data's Disparate...”, *op. cit.*, 2016, pàg. 691.

²⁰⁴WILLIAMS, B. A., BROOKS, C. F. i SHAMARGAD, Y., “How Algorithms Discriminate...”, *op. cit.*, 2018, pàg. 81.

²⁰⁵KOENE, A., “Algorithmic Bias: Addressing Growing Concerns”, *IEEE Technology and Societal Magazine*, vol. 36, núm. 2, 2017, pàg. 32.

²⁰⁶BAROCAS, S. i SELBST, A. D., “Big Data's Disparate...”, *op. cit.*, 2016, pàg. 699.

²⁰⁷ OBSERVATORI D'ÈTICA EN INTEL·LIGÈNCIA ARTIFICIAL DE CATALUNYA (OEIAC), *Intel·ligència artificial, ètica i societat...*, *op. cit.*, 2021, pàg. 36. La majoria de persones expertes en aquest àmbit (persones desenvolupadores d'intel·ligència artificial, persones que treballen en el camp de la tecnologia i ciències socials computacionals, etc...), manifesta la impossibilitat d'un sistema totalment imparcial, ja que els sistemes d'intel·ligència artificial, especialment d'aprenentatge automàtic estan dissenyats per diferenciar entre persones, imatges o documents.

conflictual, ja que esdevé molt més fàcil cobrir una discriminació intencional a través de la discriminació no intencional que es reflecteix en les dades fruit d'opressions històriques a determinats membres de grups socials determinats. Una exemplificació d'aquest fenomen es troba en què les persones legisladores podrien intencionalment esbiaixar la recollida de dades per assegurar que el tractament suggereix normes que són menys favorables per a membres de classes protegides²⁰⁸.

També cal tenir en compte amb quina finalitat s'empra el sistema algorítmic. Tal com s'explica en relació al sistema algorítmic AMS, és dubtosa la idea que aquest es vegi com una mera eina tècnica, sinó més aviat com un esforç sociotècnic que a la vegada inclou i reforça certs objectius sociopolítics²⁰⁹. A continuació, abordem l'opacitat amb la que actuen els sistemes de presa de decisió mitjançant algoritmes.

3.2. Opacitat intrínseca de les decisions automatitzades: una manera d'emparar la discriminació

L'opacitat es defineix com la qualitat d'opac, és a dir, que no deixa passar la llum, oposat a transparent²¹⁰. Els algoritmes són opacs ja que governen múltiples esferes de la vida quotidiana de manera invisible. Invisible ja que resulten ser inescrutables entre les capes de programació informàtica, és a dir, gairebé hermètics a la interpretació, correcció i millora²¹¹. Tot i que el codi font estigui escrit amb claredat i les bases de dades siguin extremament grans (però sigui possible la seva comprensió), la interacció entre els dos elements en el mecanisme de l'algorisme és allò que produeix la complexitat (i per tant,

²⁰⁸BAROCAS, S. i SELBST, A. D., "Big Data's Disparate...", *op. cit.*, 2016, pàg. 692.

²⁰⁹ALLHUTTER D., CECH F., FISCHER F., GRILL G., i MAGER A., "Algorithmic Profiling of Job...", *op. cit.*, 2020, pàg. 13.

²¹⁰Opacitat. Diccionari en línia. Enciclopèdia Catalana. 2023. Disponible a: <https://www.diccionari.cat/GDLC/opacitat>. Darrera visita el 2 de setembre de 2023.

²¹¹MONASTERIO ASTOBIZA, A., "Ètica algorítmica: Implicaciones éticas de una sociedad cada vez más dominada por algoritmos", *Dilemata*, núm. 24, 2017, pàg. 196; AÑÓN ROIG, M.J., "Desigualdades algorítmicas...", *op. cit.*, 2022, pàg. 23. L'autora estableix que la intel·ligència artificial es presenta com una eina de gestió de riscos originàriament basada en el paradigma matemàtic, la probabilitat, l'estadística i el càlcul, l'objectiu de la qual és fer front a la incertesa mitjançant el desplegament de funcions d'anticipació que permeten realitzar operacions com reconèixer patrons, predir conductes i adoptar decisions a través de l'estadística computacional, desplaçant la probabilitat clàssica basada en el principi de causalitat i substituint-la per una tipologia de raonament que recorre a correlacions múltiples i no evidents que, per tant, provoca un augment de l'opacitat.

l'opacitat)²¹². Per això als sistemes d'intel·ligència artificial sovint se'ls relaciona amb caixes negres²¹³ ja que no està clar per ningú per què un sistema pren certa decisió sobre un mateix. És així perquè en aquest moment, atesa l'evolució dels sistemes automatitzats (que pretenen replicar la intel·ligència humana) i la complexitat de tota la informació gestionada (especialment sofisticada per la interacció entre diferents capes de dades), és difícil conèixer l'origen de la decisió o elecció que ofereix l'algoritme, és a dir, quin és el motiu pel qual considera que aquella és la millor opció²¹⁴. Aquesta opacitat ocasiona que la protecció que es pugui brindar a les persones sigui molt poc efectiva i la desconeixença per part d'aquestes dels efectes discriminatoris de les eines que estan determinant certs aspectes de les seves vides. A més, la desconeixença sobre el funcionament del sistema pot socavar l'aplicació dels acords ètics establerts²¹⁵. Quan ens referim a la discriminació, l'opacitat esdevé un dels grans obstacles per identificar-la, crear mecanismes per contrarestar-la i atenuar-la. Un debat informat és impossible sense l'habilitat d'entendre quins sistemes existents s'estan emprant, com s'estan utilitzant i com aquests sistemes poden causar perjudicis no intencionats²¹⁶.

En relació amb l'apartat anterior, fer els biaixos evidents suposaria assegurar la transparència, que sembla un objectiu quimèric per moltes raons, incloent el fet que l'algoritme més acurat és aquell menys transparent. Cal fer referència també al fet que els algoritmes estan majoritàriament implementats per entitats privades que busquen maximitzar els beneficis, operant sota les mínimes obligacions de transparència²¹⁷. En referència a l'elaboració de perfils, el Supervisor de la Protecció de Dades a Europa ha puntualitzat que el problema no és l'elaboració de perfils en si, sinó la manca d'informació significativa sobre la lògica de l'algoritme que desenvolupa aquests perfils i té un efecte en la persona que cedeix les seves dades²¹⁸. En conseqüència, l'opacitat de

²¹² BURRELL J., "How the machine 'thinks': Understanding opacity in machine learning algorithms", *Big Data & Society*, vol. 3, núm. 1, 2016, pàg. 5.

²¹³ Terme denominat en anglès "black boxes". En referència amb la primera al·lusió a aquest terme, vegeu: PASQUALE, F., *The Black box society: The secret algorithms that control Money and information*, Cambridge (Massachusetts), Harvard University Press, 2015.

²¹⁴ RIVAS VALLEJO, P., "Discriminación algorítmica: detección, prevención y tutela", *op. cit.*, 2021, pàg. 8.

²¹⁵ HEINRICH, B., "Discrimination in the age of artificial intelligence", *AI & SOCIETY*, vol. 37, núm. 1, 2022, pàg. 151.

²¹⁶ REISMAN, D., SCHULTZ, J., CRAWFORD, K. i WHITTAKER, M., *Algorithmic Impact Assessments: a Practical Framework for Public Agency Accountability*, AI Now Institute, New York, April 2018, pàg. 3.

²¹⁷ NOTO LA DIEGA, G., "Against the Dehumanisation...", *op. cit.*, 2018, pàg. 9.

²¹⁸ EUROPEAN DATA PROTECTION SUPERVISOR, *Opinion 3/2015 (with addendum)*, *Europe's big opportunity. EDPS recommendations on the EU's options for data protection reform*, Brussels, 9 October 2015, pàg. 8.

les decisions automatitzades mitjançant sistemes algorítmics és un obstacle majúscul per fer front a la discriminació primordialment per dues raons. En primer lloc, perquè moltes vegades a causa de la seva complexitat és impossible saber quin és el seu funcionament i el procés de decisió que segueixen, impedit la comprensió de com operen, a vegades inclús per part d'aquelles persones que l'han creat, dificultant determinar perquè són discriminatoris; i, en segon lloc, perquè ocasiona que les persones que són discriminades ho desconeguin i no puguin actuar al respecte.

La subtilitat d'aquest tipus de discriminació suposa una dificultat també per saber on s'ha de mirar per obtenir evidències rellevants que revelin discriminació *prima facie*. Les persones discriminen en base a actituds negatives (estereotips, prejudicis), i biaixos no intencionals (pràctiques organitzacionals o estereotips internalitzats) que poden actuar com una alerta que la discriminació s'ha dut a terme; en canvi, no existeixen mecanismes equivalents en els sistemes de decisió automatitzada²¹⁹. No es pot assumir que els sistemes automatitzats discriminaran de manera similar les persones, o empraran patrons familiars de discriminació. Justament el que queda més fora de l'abast de la supervisió són les conclusions a les que arriba el sistema algorítmic amb les dades que té, que poden no ser raonables.

La regulació antidiscriminació està dissenyada per prevenir formes conegudes de prejudici humà i discriminació, tals com la desigualtat de gènere en el lloc de treball o el racisme en les pràctiques de contractació. Les noves formes de discriminació algorítmica són cada cop més difícils de detectar i poden no tenir cap base per enfrontar-les ni en la llei ni en la jurisprudència. Sabem que la retribució pot actuar com a dada rellevant per establir el gènere d'una persona, però com poden correlacionar-se amb atributs protegits el comportament de la navegació a la xarxa o els clics en pàgines web?²²⁰. Existeix una pèrdua de la traçabilitat, de manera que aquests sistemes esdevenen inaccessibles a la ciutadania i cada cop més difícils de comprendre o desafiar. Addicionalment, les pràctiques contemporànies de recollida i tractament de dades interactuen amb la persona sota unes bases invisibles d'automatització, a la qual cosa s'afegeix l'assumpció que un

²¹⁹ALLEN, R. i MASTERS, D., "Artificial Intelligence: the right to protection from discrimination caused by algorithms, machine learning and automated decision-making", *ERA Forum*, vol 20, núm. 4, 2020, pàg. 597. S'afirma en l'article que no existeix un mecanisme creïble segons el qual les persones puguin entendre com les decisions s'han pres sobre elles.

²²⁰WACHTER, S., MITTELSTADT, B. i RUSSELL, C., "Why fairness cannot be automated: Bridging the gap between EU non-discrimination law and AI", *Computer Law and Security Review*, vol 41, 2021, pàg. 66 i 67.

gran percentatge d'aquestes dades s'ha obtingut per part del sector privat amb finalitats empresarials²²¹. En moltes ocasions, també, el dret a l'accés a la justícia es veu condicionat per les pròpies limitacions amb que es troben les persones derivades de les despeses econòmiques, burocràtiques i administratives, la por a una violació de la privacitat, a represàlies, a una restricció o pèrdua de drets adquirits que són necessaris per al desenvolupament en la vida quotidiana de la persona en qüestió²²², o pel fet que sovint, les víctimes de discriminació no són conscients que els seus drets estan essent vulnerats²²³. Una persona descartada en un procés de selecció de personal dut a terme per un algoritme, que ha basat la seva decisió en una categoria sensible, difícilment serà conscient de les raons per les quals no ha estat contractada. Serà especialment complex saber-ho si no té accés o coneixement de l'impacte dels resultats de l'algoritme en altres persones que hagin participat en el mateix procés²²⁴. Així i tot, hi ha casos en els quals el grau de personalització al qual poden arribar els sistemes algorítmics pot ocasionar inclús la dificultat de trobar una comparació vàlida entre persones que s'hagin vist sotmeses a un mateix sistema algorítmic.

3.3. Personalització en els sistemes algorítmics

Els sistemes algorítmics tendeixen a la personalització dels casos. El resultat d'un algoritme es basarà en una combinació de característiques i comportaments únics d'una persona (en el cas de l'elaboració d'un perfil d'una persona) o grup de persones (en el cas de l'elaboració d'un perfil de grup)²²⁵. El nivell de detall al que poden arribar els sistemes algorítmics per dur a terme classificacions no té precedents.

²²¹LEESE, M., "The New Profiling: Algorithms, Black Boxes...", *op. cit.*, 2014, pàg. 504.

²²² ORGANISATION FOR SECURITY AND CO-OPERATION IN EUROPE (OSCE), *Use of actio popularis in cases of discrimination*, Skopje, November 2016, pàg. 16.

²²³KATYAL, S. K., "Private Accountability in the Age of Artificial Intelligence", *U.C.L.A. Law Review*, vol. 66, 2019, pàg. 100.

²²⁴ SORIANO ARNANZ, A., "Decisiones automatizadas y discriminación...", *op. cit.*, 2021, pàg. 33.

²²⁵ COTINO HUESO, L., "Nuevo paradigma en la garantías de los derechos fundamentales y una nueva protección de datos frente al impacto social y colectivo de la inteligencia artificial", a COTINO HUESO, L., (dir.), *Derechos y garantías ante la inteligencia artificial y las decisiones automatizadas*, Navarra, Aranzadi, 2022, pàg. 87-88. L'autor explica que la identitat dels individus que conformen el grup no és rellevant, encara que es percebi com un procediment tècnic d'individualització, la personalització no implica individualització, de manera que la majoria de persones no estan perfilades individualment sinó com a membres d'un grup específic que es pot definir segons el tipus d'adquisicions, barri, afinitats

Per exemple, s'ha esmentat que el sistema algorítmic AMS té una forta tendència a individualitzar les causes de desocupació²²⁶. Això suposa que com l'algoritme té accés a tantes dades, pot crear respostes individuals per a cada persona que analitza, de manera que no només serà difícil trobar altres amb qui comparar-se sinó que també serà difícil trobar elements comuns que hagin dut l'algoritme a prendre una decisió discriminatòria. De la mateixa manera, les botigues de venda en línia poden diferenciar el preu per productes idèntics basant-se en la informació que la botiga té sobre la persona consumidora. Una botiga pot reconèixer qui visita la seva web mitjançant galetes i categoritzar-los com a sensibles al preu o no sensibles al preu. Amb aquesta pràctica es pretén cobrar a cada persona consumidora el màxim preu que estigui disposada a pagar²²⁷.

Aquest fenomen porta a una altra problemàtica relativa als factors que han determinat la decisió i com la determinació d'aquests factors pot tenir resultats discriminatoris. A Estats Units d'Amèrica del Nord, un estudi fet sobre la fixació dels preus en les assegurances de cotxe va revelar que les primes es calculaven en base a les qualificacions creditícies de les persones assegurades, en detriment d'altres factors clarament influents pel risc com l'historial de la persona com a conductora. Les empreses asseguradores accedien als informes de solvència de les persones i escollien diverses dades a les quals aplicaven els algorismes per fixar les primes. Donada la rellevància que adquiria la dada sobre la solvència de la persona conductora, com més solvent era, més baixa era la prima (atès que hi havia menys risc d'impagament), de manera que es demanaven primes més altes als sectors més vulnerables econòmicament, empitjorant la seva situació econòmica. També suposava que la forma de tarifar era completament oculta per les persones assegurades que no sabien la rellevància que podia tenir la seva qualificació creditícia²²⁸. El Joint Committee de las European Supervisory Authorities (ESAs) va destacar el risc de la falta de transparència en la fixació de preus, ja que com més s'individualitzava el risc, més difícil era per a les persones consumidores saber si el preu que oferia una

geogràfiques, etc. Inclús, aquests grups perfilats es poden alimentar amb dades agregades, que són aquelles no vinculades a persones (open data, IoT, geolocalització, navegació, etc.).

²²⁶ALLHUTTER, D., CECH, F., FISCHER, F., GRILL, G. i MAGER, A., "Algorithmic Profiling of Job...", *op. cit.*, 2020, pàg. 9.

²²⁷ZUIDERVEEN, F. J. i POORT, J., "Online price discrimination and EU data privacy law", *Journal of Consumer Policy*, vol. 40, 2017, pàg. 351.

²²⁸MUÑOZ PAREDES, M.L., "«Big data» y contrato de seguro: los datos generados por los asegurados y su utilización por los aseguradores", a HUERGO LORA, A.J., (dir.), i DÍAZ GONZÁLEZ, G. M. (coord.), *La regulación de los algoritmos*, *op. cit.*, 2020, pàg. 148-151.

asseguradora era competitiu, atès que no podria fer una tasca comparativa per descobrir-ho²²⁹.

La personalització que proporcionen els algoritmes genera que puguin no haver-hi dos casos iguals. Una persona que s'ha vist discriminada hauria de buscar el factor comú (per exemple, l'origen ètnic) pel qual ha estat discriminada i altres persones que s'han vist discriminades pel mateix motiu. Tanmateix, aquestes altres persones poden haver estat discriminades per altres factors que no comparteixen amb la primera persona, atomitzant cada cop més els grups de persones que es veuen perjudicades. Pot no haver només un criteri en base al qual es pot establir que un tracte diferencial ha estat discriminatori, sinó que hi ha molts criteris susceptibles d'haver causat el resultat discriminatori. Com veurem al llarg d'aquest treball, aquest és un dels factors que ha causat que s'hagi recorregut a reclamar una perspectiva interseccional de la discriminació, és a dir, l'estudi de la manera per la qual els diferents tipus de discriminació es relacionen o s'afecten entre si²³⁰, i que s'hagi creat una categorització concreta per aquest tipus de casos anomenada discriminació interseccional²³¹.

Finalment, un altre problema que podem identificar, i que és un risc afegit a la lluita envers la discriminació, és la magnificació dels efectes d'una decisió presa mitjançant sistemes algorítmics. Al contenir moltes més dades, es posseeix molta més informació sobre moltes més persones i la repercussió de les decisions és exponencialment més gran, ocasionant que arran d'una decisió, moltes més persones puguin ser discriminades.

3.4. Magnificació dels efectes de la discriminació

Habitualment, la prohibició de discriminació sol entrar en conflicte amb l'autonomia de la voluntat. Els factors que s'han de tenir en compte quan es fa una ponderació entre ells

²²⁹EUROPEAN SUPERVISION AUTHORITIES (ESAs), *Joint Committee Final Report on Big Data*, JC/2018/04, Brussels, 15 March 2018, pàg. 12.

²³⁰Intersectionality. Cambridge Dictionary. Cambridge University Press. 2023. Disponible a: <https://dictionary.cambridge.org/es/diccionario/ingles/intersectionality>. Darrera visita el 2 de setembre de 2023.

²³¹ La Ley 15/2022, de 12 de julio, integral para la igualdad de trato y la no discriminación defineix discriminació interseccional com aquella que es produeix quan concorren o interactuen diverses causes de les previstes en la llei, generant una forma específica de discriminació.

són la repercussió social de la discriminació, la posició dominant de l'entitat discriminadora en el mercat, el grau d'afectació a la dignitat²³² i a la integritat moral de la víctima o les víctimes en el cas concret²³³.

Aquí és on podem localitzar un problema majúscul que planteja la presa de decisions mitjançant algorismes, i és que la repercussió social de la discriminació que ocasiona -és a dir els seus efectes-, és molt àmplia. Mentre que una persona pot discriminar potencialment contra centenars de persones, un algoritme -amb el seu il·limitat rendiment, té la capacitat per discriminar contra milers de persones²³⁴. Per tant, els processos algorítmics de presa de decisions no només poden frustrar els objectius de les lleis antidiscriminatòries, sinó servir per reproduir desigualtats a major escala²³⁵ (inclús al marge de qualsevol intenció de discriminar en relació a alguna característica de la persona, ja que aquesta pot estar estadísticament associada a un determinat criteri²³⁶).

Noble examina, per exemple, les maneres en les quals els resultats dels motors de recerca poden perpetuar narratives que reflecteixen distribucions de poder desiguals històriques en la societat²³⁷. Aquesta amplitud en el seu àmbit d'incidència fa que siguin mecanismes proclius a ocasionar discriminació sistèmica, que es pot definir com aquella lligada a polítiques que afecten a moltes persones i a pràctiques generalitzades que provoquen l'efecte combinat de perjudicar sistemàticament a grups vulnerables i fer-ho de manera permanent²³⁸. D'aquesta manera, mitjançant l'anàlisi massiva de dades, és lògic plantejar que la discriminació afecta sempre a més d'una persona ja que els mètodes deductius sempre són els mateixos, i l'afectació a més d'un cas desemboca en segregació, és a dir, discriminació grupal resultat natural d'una discriminació sistemàtica i estructural que pot ser social, laboral o educativa, entre d'altres, i que es caracteritza per estar dirigida a tot un grup en desavantatge i no a una persona integrant del grup en desavantatge (abordem

²³² Vegeu, en aquest sentit: KAMINSKI, M.E., "Binary Governance...", *op. cit.*, 2019, pàg. 1543-1544.

²³³ BILBAO, J.M., "Prohibición de discriminación y las relaciones entre particulares", *Teoría y realidad constitucional*, núm. 18, 2006, pàg. 155-157.

²³⁴ BIGMAN, Y. E., WILSON, D., ARNESTAD, M. N. WAYTZ, A. i GRAY, K., "Algorithmic Discrimination Causes Less Moral Outrage Than Human Discrimination", *Journal of Experimental Psychology: General*, vol. 152, núm. 1, 2023, pàg. 5 i 6.

²³⁵ EUBANKS, V., *Automating Inequality: How high-Tech Tools Profile, Police, and Punish the Poor*, New York, St. Martin's Press, 2018, pàg. 190. L'autora especifica que: "we should all care about the digital poorhouse because it intensifies discrimination".

²³⁶ SÁEZ LARA, C., "El algoritmo como protagonista de la relación laboral...", *op. cit.*, 2020, pàg. 45-46.

²³⁷ NOBLE, S. U. *Algorithms of Oppression: How Search Engines Reinforce Racism*, *op. cit.*, 2018, pàg. 71.

²³⁸ MERCAT-BURNS, M., "Systemic discrimination...", *op. cit.*, 2018, pàg. 1.

el desafiament col·lectiu de decisions automatitzades discriminatòries en el capítol VII d'aquest treball).

4. Conclusions parcials

En aquest capítol plantejàvem la qüestió relativa a si les decisions automatitzades augmenten les possibilitats de discriminació i, si és així, per què.

A mode introductori, hem fet una presentació de certs elements clau per aquest treball, la igualtat, l'equitat i la discriminació, i els hem relacionat amb els sistemes algorítmics. Seguidament hem plantejat com funciona la discriminació algorítmica i quines són les formes amb les quals es posa en pràctica, ja sigui en base a mesures tècniques de funcionament dels sistemes algorítmics o en base a apreciacions humanes. Finalment, havent plantejat com es duu a terme, hem reflexionat sobre els riscos afegits que es poden identificar per a la lluita contra la discriminació arran de l'existència de decisions automatitzades mitjançant sistemes algorítmics, concretant-se en la creació de biaixos algorítmics per part de la pròpia inèrcia del sistema o per part de les persones que actuen en totes les fases del seu funcionament, en l'opacitat intrínseca al sistema, en el grau de personalització al qual pot arribar el sistema, i en la magnificació dels efectes discriminatoris causats per la capacitat d'abastar una quantitat molt gran (massiva) de dades per prendre les decisions, que tenen el potencial d'afectar moltes persones.

Per consegüent podem concloure que les decisions automatitzades augmenten les possibilitats de discriminació principalment perquè afegeixen certs riscos a la forma tradicional de discriminació.

A més, els sistemes algorítmics es van presentar inicialment com una eina més objectiva, i per tant més respectuosa amb la igualtat, que la presa de decisions duta a terme per éssers humans, ja que s'entenia que era una màquina la que decidia, inherentment objectiva, i no una persona, inherentment subjectiva. Els estudis que han evidenciat que les màquines podien contenir els mateixos biaixos i prejudicis que les persones que les havien ideat han participat a posar en dubte aquest discurs. No obstant això, sota la nostra perspectiva, la normativa antidiscriminació parteix en desavantatge per fer front a un gran producte que

pot arribar (o ja ho està fent) a produir grans beneficis econòmics, com és la intel·ligència artificial, i aquesta qüestió és d'especial rellevància en l'estudi sobre eines del dret envers la discriminació algorítmica.

Després de concretar les dificultats per la lluita contra la no discriminació que ocasionen les decisions automatitzades mitjançant els sistemes algorítmics, estudiem en el següent capítol els límits de la regulació europea antidiscriminació per abordar les casuístiques específiques de la discriminació algorítmica.

Capítol II. Els límits de la regulació de la discriminació algorítmica en la normativa europea antidiscriminació

Havent analitzat en el capítol I, com i perquè les decisions automatitzades mitjançant sistemes algorítmics poden causar discriminació, abordem en aquest capítol i en el següent la segona qüestió de recerca, a saber, si el marc regulador europeu antidiscriminatori i el marc regulador europeu de protecció de dades són suficients per fer front a les discriminacions arran de decisions automatitzades mitjançant sistemes algorítmics. Per respondre a aquesta pregunta, en aquest capítol II del treball presentem els punts febles que presenta el marc regulador europeu antidiscriminatori i si aquests ocasionen que la normativa no sigui suficient per abordar la discriminació algorítmica. Hem organitzat el raonament en base a les quatre problemàtiques que considerem que poden tenir més rellevància per la discriminació algorítmica, a saber, l'organització del dret antidiscriminatori en base a categories protegides i les diferències de protecció que brinda segons la categoria i l'àmbit d'aplicació (1), el fet que es tingui molt poc en compte l'element col·lectiu de la discriminació, clarament definitori de la discriminació algorítmica (2), les deficiències de la conceptualització bidimensional de discriminació directa i indirecta (3), i les problemàtiques que presenten les exigències probatòries actuals en casos en què es produeix discriminació algorítmica (4). Finalment, resumim els elements tractats en aquest capítol en un apartat de conclusions parcials (5).

1. Categories protegides i àmbits d'aplicació material

En aquest apartat abordem els límits que sorgeixen com a conseqüència que la normativa antidiscriminació europea s'articuli en base a certs atributs protegits. Primerament expliquem la raó de l'existència de la normativa en matèria antidiscriminació a la UE (1.1), per abordar seguidament les conseqüències en la discriminació algorítmica d'una protecció basada excessivament en certs atributs protegits (1.2). En dret derivat la protecció d'aquests atributs protegits es concreta en les directives antidiscriminació

europées, que proporcionen una protecció diferent segons l'atribut protegit generant un “*abast asimètric de la legislació antidiscriminatòria a la UE*”²³⁹(1.3) .

1.1. Introducció al dret antidiscriminatori europeu

Va ser en finalitzar la Segona Guerra Mundial, i internacionalitzar-se els Drets Humans posteriorment, gràcies a l'activitat i desenvolupament d'organitzacions internacionals, que es van reconèixer en els règims jurídics estatals els drets relatius a la igualtat i a la no discriminació. L'inici del reconeixement troba la seva fonamentació tant en les diferents experiències socials sorgides al final de la segona guerra mundial com en les lluites dels moviments socials per aconseguir drets civils igualitaris²⁴⁰. És llavors quan es van començar a crear normes de dret positiu per millorar la igualtat de tracte entre les persones respecte les característiques causants de la marginació²⁴¹.

Els nous drets, i sobretot el dret a la igualtat, es troba especialment globalitzat, essent aquest un dels seus trets definitoris²⁴². Es pot considerar que la prohibició de discriminar en sentit ampli recull el dret a la igualtat i, en canvi, la prohibició de discriminar en sentit estricte recull el dret antidiscriminatori²⁴³. El dret clàssic a la igualtat juga enfront l'Estat i en canvi el dret antidiscriminatori ho fa també, i potser, sobretot, enfront altres particulars.

Les primeres referències normatives sobre discriminació es troben en el Tractat Constitutiu de la Comunitat Europea (CEE), on s'establí en els articles 12 i 119 la prohibició de discriminació per motiu de nacionalitat i igualtat salarial entre sexes, respectivament. Cal subratllar que a l'inici de la construcció de la protecció envers la discriminació a la UE, l'àmbit tutelat enfront el tracte discriminatori es basava en les

²³⁹ GELLERT, R. *et al.*, “A comparative analysis of anti-discrimination and data protection legislations”, a CUSTERS, B., CALDERS, T., SCHERMER, B i ZARSKY, T., (eds.), *Discrimination and privacy in the information society: data mining and profiling in large databases*, Berlin, Springer, 2013, pàg. 63-66.

²⁴⁰ Com a conseqüència que el dret antidiscriminatori estigui tan estretament lligat a la doctrina internacional de drets humans, totes les jurisdiccions treballen sobre una significació de “discriminació” comunament acordada.

²⁴¹ PUIG HERNÁNDEZ, M. A., *La protección de la libertad ideológica frente al trato discriminatorio en la relación jurídico-laboral*, Tesi doctoral, Universitat de Barcelona, Barcelona, 2018, pàg. 21-22.

²⁴² REY MARTÍNEZ, F., *Derecho Antidiscriminatorio*, *op. cit.*, 2019, pàg. 22.

²⁴³ *Ídem*, pàg. 47.

relacions econòmiques sociolaborals i, a més, per raó de sexe, és a dir, la relació de les normes antidiscriminatòries amb l'àmbit social es donava per les diferències de tracte entre homes i dones al treball, segons el context econòmic europeu. De manera que aquestes normes van aparèixer com un mecanisme auxiliar i complementari del principi d'integració comunitari de caràcter marcadament econòmic²⁴⁴.

A partir de la sentència Defrenne (II)²⁴⁵, la prohibició de discriminació va quedar establerta com un principi general del dret comunitari i així va quedar reflectida posteriorment en l'article 13 del Tractat d'Amsterdam de 1997²⁴⁶. Aquest tractat va donar un gran impuls a la legislació antidiscriminació creant, des del punt de vista institucional nombrosos organismes nacionals especialitzats i dedicats a promoure la igualtat de tracte reconstruint també organismes ja existents²⁴⁷. També, aquest mateix Tractat va expandir significativament la competència de la UE per combatre la discriminació que originalment estava reduïda a la discriminació per raó de sexe, ja que l'article 13²⁴⁸ va introduir com a motius objecte de discriminació les creences, l'origen ètnic, la discapacitat, l'edat i l'orientació sexual (aquest article actualment és el 19 del TFUE, que habilita a prendre accions per combatre la discriminació per raó de sexe, raça o origen ètnic, religió o creença, discapacitat, edat o orientació sexual).

L'Estratègia Europea de Treball i la convergència de les tradicions jurídiques nacionals sobre el referent normatiu comunitari van ser causants de la promulgació de la Directiva 2000/43/CE del Consejo, de 29 de junio de 2000, relativa a la aplicación del principio de igualdad de trato de las personas independientemente de su origen racial o étnico²⁴⁹ (d'ara en endavant, Directiva 2000/43/CE); la Directiva 2000/78/CE del Consejo, de 27 de

²⁴⁴PUIG HERNÁNDEZ, M. A., *La protección de la libertad ideológica...*, op. cit., 2018, pàg. 22; GELLERT, R. et al., "A comparative analysis...", op. cit., 2013, pàg. 67-68. En aquest article s'explica que les lleis antidiscriminació en aquell moment no es concebien com a drets fonamentals *per se*, sinó com a eines per facilitar la mobilitat i el funcionament del mercat intern: combatre la discriminació entre la ciutadania de la UE basada en la nacionalitat i el sexe en assumptes relacionats amb el treball eren maneres de millorar l'eficiència del mercat comú i prevenir la discriminació per motius que eren econòmicament ineficients.

²⁴⁵TJCE, sentència de 8 d'abril de 1976, Assumpte 43-75, Gabrielle Defrenne contra Société anonyme belge de navigation aérienne Sabena.

²⁴⁶PUIG HERNÁNDEZ, M. A., *La protección de la libertad ideológica...*, op. cit., 2018, pàg. 120.

²⁴⁷RORIVE, I., "A comparative and European examination of National Institutions in the field of racism and discrimination", a BOYLE, K., (ed.), *New Institutions for human Rights Protection*, Oxford, Oxford University Press, 2009, pàg. 138 i s.

²⁴⁸Article 13 TCE.

²⁴⁹ Directiva 2000/43/CE del Consejo, de 29 de junio de 2000, relativa a la aplicación del principio de igualdad de trato de las personas independientemente de su origen racial o étnico. DOCE, L 180/22, 19.07.2000.

noviembre de 2000, relativa al establecimiento de un marco general para la igualdad de trato en el empleo y la ocupación²⁵⁰ (d'ara en endavant, Directiva 2000/78/CE)²⁵¹; com també posteriorment la Directiva del Consejo 2004/113/CE, de 13 de diciembre de 2004, por la que se aplica el principio de igualdad de trato entre hombres y mujeres al acceso a bienes y servicios y su suministro²⁵² (d'ara en endavant, Directiva 2004/113/CE) i; la Directiva 2006/54/CE del Parlamento Europeo y del Consejo, de 5 de julio de 2006, relativa a la aplicación del principio de igualdad de oportunidades e igualdad de trato entre hombres y mujeres en asuntos de empleo y ocupación (refundición)²⁵³ (d'ara en endavant, Directiva 2006/54/CE), que, com ja hem esmentat, conjuntament anomenarem directives antidiscriminació.

Actualment, la no discriminació en la normativa de la Unió Europea és un principi general consagrat en l'article 2²⁵⁴ i l'article 3.3 del Tractat de la Unió Europea (d'ara en endavant, TUE)²⁵⁵. L'article 10 del TFUE integra la lluita envers la discriminació per raó de sexe, raça o origen ètnic, religió o creences, discapacitat, edat o orientació sexual en les polítiques i activitats de la UE. L'article 18 del TFUE²⁵⁶ prohibeix qualsevol discriminació basada en la nacionalitat (l'única característica protegida en tots els àmbits del dret de la UE²⁵⁷), i l'article 19 del TFUE²⁵⁸ que habilita a dur a terme accions per lluitar envers la discriminació per motiu de sexe, origen racial o ètnic, religió o creences, discapacitat, edat o orientació sexual (i és l'article del qual dimanen les directives antidiscriminació).

En l'àmbit de la UE cal destacar també la Carta dels Drets Fonamentals de la Unió Europea (d'ara en endavant, CDFUE), de l'any 2000. En el seu article 21, es prohibeix tota discriminació, i en particular, l'exercida per raó de sexe, la raça o el color, orígens

²⁵⁰ Directiva 2000/78/CE del Consejo, de 27 de noviembre de 2000, relativa al establecimiento de un marco general para la igualdad de trato en el empleo y la ocupación. DOCE, L 303/16, 2.12.2002.

²⁵¹BELL, M., *Anti-Discrimination Law and the European Union. Oxford Studies in European Law*, Oxford, Oxford University Press, 2002, pàg. 205-212.

²⁵² Directiva del Consejo 2004/113/CE, de 13 de diciembre de 2004, por la que se aplica el principio de igualdad de trato entre hombres y mujeres al acceso a bienes y servicios y su suministro. DOUE, L 373/37, 21.12.2004.

²⁵³ Directiva 2006/54/CE del Parlamento Europeo y del Consejo, de 5 de julio de 2006, relativa a la aplicación del principio de igualdad de oportunidades e igualdad de trato entre hombres y mujeres en asuntos de empleo y ocupación (refundición). DOUE, L 204/23, 26.7.2006.

²⁵⁴Article 2 del TUE.

²⁵⁵Article 3.3 del TUE.

²⁵⁶Article 18 del TFUE.

²⁵⁷ TOBLER, C., *Indirect discrimination under Directives 2000/43 and 2000/78*, European network of legal experts in gender equality and non-discrimination, Brussels, 2022, pàg. 92.

²⁵⁸ Article 19 del TFUE.

ètnics o socials, característiques genètiques, llengua, religió o conviccions, opinions polítiques o de qualsevol altre tipus, pertinença a una minoria nacional, patrimoni, naixement, discapacitat, edat o orientació sexual²⁵⁹. A l'inici, únicament tenia caràcter declaratiu. Amb l'entrada en vigor de l'última de les reformes del TUE a través del tractat de Lisboa el desembre de 2009, la CDFUE té el mateix valor jurídic que els tractats, com enuncia l'article 6 TUE²⁶⁰, considerant-se dret originari²⁶¹. Les persones poden reclamar si consideren que la legislació de la UE o la legislació estatal en aplicació de la primera no respecta la CDFUE. També es pot sol·licitar orientació sobre la seva interpretació a través del procediment de qüestió prejudicial previst en l'article 267 TFUE²⁶².

En l'àmbit del Consell d'Europa (Organització Internacional de cooperació de la qual tots els estats membres de la UE en són membres) trobem compromisos per garantir la protecció dels drets humans, principalment en el Conveni Europeu de Drets Humans (d'ara en endavant, CEDH), adoptat pel Consell d'Europa el 1950 i que va entrar en vigor el 1953, -signat per tots els estats membres de la UE-²⁶³.

L'article 14 del CEDH garanteix la igualtat de tracte en el gaudi dels drets i llibertats reconeguts en el Conveni²⁶⁴. Amb el Protocol 12 al CEDH²⁶⁵, aprovat pel Comitè de Ministres del Consell d'Europa el 26 de juny del 2000, es va ampliar l'abast de la prohibició de discriminació, en garantir la igualtat de tracte en el gaudi de tots els drets (inclosos aquells derivats de la legislació nacional). Nogensmenys, les obligacions recollides en el Protocol 12 només generen deures pels estats que el firmen, que en l'actualitat no són tots els estats membres de la UE²⁶⁶. Per consegüent és important

²⁵⁹ Article 21.1 CDFUE. En el segon apartat de l'article "*es prohibeix tota discriminació per raó de nacionalitat en l'àmbit d'aplicació del Tractat constitutiu de la Comunitat Europea i del Tractat de la Unió Europea i sense perjudici de les disposicions particulars d'aquests Tractats*".

²⁶⁰ Article 6.1 del TUE.

²⁶¹ ARCOS VARGAS, M., "La no discriminación en el derecho derivado de la Unión Europea", a MORALES ORTEGA, J. M., (coord.), *Realidad social y discriminación: Estudios sobre diversidad e inclusión laboral*, Murcia, Laborum, 2022, pàg. 21.

²⁶² Article 267 del TFUE.

²⁶³ L'article 6.2 del TUE recull l'obligació de la UE d'adherir-se al CEDH. Actualment la UE encara no s'ha adherit al CEDH.

²⁶⁴ Transcrivim aquí l'article 14 del CEDH: "*El gaudiment dels drets i les llibertats reconeguts en el present Conveni ha de ser assegurat sense cap distinció, especialment per raons de sexe, raça, color, llengua, religió, opinions polítiques o altres, origen nacional o social, pertinença a una minoria nacional, fortuna, naixença o qualsevol altra situació.*"

²⁶⁵ Transcrivim aquí l'article 1 del Protocol 12 del CEDH: "*El gaudiment de tot dret establert per la llei ha de ser assegurat, sense cap discriminació, basada en particular en el sexe, la raça, el color, la llengua, la religió, les opinions polítiques o qualsevol altra opinió, l'origen nacional o social, la pertinença a una minoria nacional, la fortuna, el naixement o qualsevol altra situació.*"

²⁶⁶ ARCOS VARGAS, M., "La no discriminación en el derecho derivado...", *op. cit.*, 2022, pàg. 18 i 19.

destacar que l'article 14 del CEDH prohibeix la discriminació únicament en relació amb l'exercici d'altres drets reconeguts en el Conveni i només els estats membres que han ratificat el Protocol 12 confereixen caràcter independent a la prohibició de discriminació sense que estigui vinculada a un altre dret del Conveni²⁶⁷.

Per posar en context la regulació antidiscriminació en l'àmbit espanyol, cal fer esment a l'article 14 de la Constitució Espanyola, on s'estableix que els espanyols són iguals davant la llei, sense que pugui prevaler cap discriminació per raó de naixença, raça, sexe, religió, opinió o qualsevol altra condició o circumstància personal o social²⁶⁸. Les directives antidiscriminació s'han transposat a Espanya a través de la Ley 62/2003, de 30 de diciembre, de medidas fiscales, administrativas y del orden social²⁶⁹ (d'ara en endavant, Llei 62/2003), i de la Ley Orgànica 3/2007, de 22 de marzo, para la igualdad efectiva de mujeres y hombres²⁷⁰ (d'ara en endavant, Llei Orgànica 3/2007), incorporant d'aquesta manera els elements i salvaguardes del dret antidiscriminatori desenvolupats i avançats per la UE²⁷¹. També cal destacar que l'any 2022 es va aprovar la Ley 15/2022, de 12 de julio, integral para la igualdad de trato y la no discriminación ²⁷² (d'ara en endavant, Llei 15/2022), que dedica l'article 23 a la intel·ligència artificial i les decisions automatitzades²⁷³, i especifica en l'article 3.1.o que la llei s'aplicarà a la intel·ligència

²⁶⁷ *Ídem*, pàg. 19.

²⁶⁸ Article 14 de la Constitució Espanyola.

²⁶⁹ Ley 62/2003, de 30 de diciembre, de medidas fiscales, administrativas y del orden social. BOE, núm. 313, de 31 de diciembre de 2003. España.

²⁷⁰ Ley Orgànica 3/2007, de 22 de marzo, para la igualdad efectiva de mujeres y hombres. BOE, núm. 71 de 23 de marzo de 2007. España.

²⁷¹ EGUÍLUZ CASTAÑEIRA, J.A., "Desafíos y retos que plantean las decisiones automatizadas y los perfilados para los derechos fundamentales", *Estudios de Deusto*, vol. 68, núm. 2, 2020, pàg. 336.

²⁷² Ley 15/2022, de 12 de julio, integral para la igualdad de trato y la no discriminación. BOE, núm. 167, de 13 de julio de 2022. España. També, l'any 2021 es va aprovar la Carta de Derechos Digitales, que dedica el segon capítol a drets d'igualtat relacionats amb l'entorn digital, però com s'especifica en la mateixa carta, no té caràcter normatiu.

²⁷³ L'article 23 de la Llei 15/2022 estableix: "1. En el marco de la Estrategia Nacional de Inteligencia Artificial, de la Carta de Derechos Digitales y de las iniciativas europeas en torno a la Inteligencia Artificial, las administraciones públicas favorecerán la puesta en marcha de mecanismos para que los algoritmos involucrados en la toma de decisiones que se utilicen en las administraciones públicas tengan en cuenta criterios de minimización de sesgos, transparencia y rendición de cuentas, siempre que sea factible técnicamente. En estos mecanismos se incluirán su diseño y datos de entrenamiento, y abordarán su potencial impacto discriminatorio. Para lograr este fin, se promoverá la realización de evaluaciones de impacto que determinen el posible sesgo discriminatorio. 2. Las administraciones públicas, en el marco de sus competencias en el ámbito de los algoritmos involucrados en procesos de toma de decisiones, priorizarán la transparencia en el diseño y la implementación y la capacidad de interpretación de las decisiones adoptadas por los mismos. 3. Las administraciones públicas y las empresas promoverán el uso de una Inteligencia Artificial ética, confiable y respetuosa con los derechos fundamentales, siguiendo especialmente las recomendaciones de la Unión Europea en este sentido. 4. Se promoverá un sello de calidad de los algoritmos". Per una visió crítica amb el contigut d'aquest article, vegeu: VICENTE PALACIO, A., "La inteligencia artificial en la ley integral para la igualdad y no discriminación", *Revista*

artificial, gestió massiva de dades, així com altres esferes d'anàloga significació²⁷⁴. A escala autonòmica, a Catalunya, l'any 2020, es va aprovar la Llei 19/2020, del 30 de desembre, d'igualtat de tracte i no discriminació²⁷⁵ (d'ara en endavant, Llei 19/2020).

1.2. Protecció en base a certs atributs protegits

L'article 19 del TFUE, en el seu primer apartat, disposa que, sense perjudici d'altres disposicions dels tractats i dins els límits de les competències atribuïdes a la UE per aquests, el Consell per unanimitat, en base al procediment legislatiu especial i prèvia aprovació del Parlament Europeu podrà adoptar accions adequades per lluitar contra la discriminació per motiu de sexe, origen racial o ètnic, religió o creences, discapacitat, edat o orientació sexual²⁷⁶. En el segon apartat, es precisa que el Parlament Europeu i el Consell podran adoptar principis bàsics de les mesures de la UE d'estímul per donar suport a les mesures empreses pels estats membres, excloent tota harmonització de les disposicions legals i reglamentàries dels estats membres²⁷⁷.

La tipologia de discriminacions produïdes pels algorismes (que comporta l'anàlisi de dades personals i de comportament) augmenta el risc de discriminació múltiple (que el TJUE encara no ha reconegut), és a dir, aquella discriminació basada en més d'un atribut protegit²⁷⁸ -o també, en base a algun atribut no estrictament protegit. Donada l'existència d'indicadors i correlacions estadístiques, es canvia la naturalesa de la discriminació augmentant les possibilitats de mala categorització de les dades de comportament, essent

General de Derecho del Trabajo y de la Seguridad Social, núm. 64, 2023, pàg. 127. L'autora afirma que el contingut normatiu de l'article 23 de la Llei 15/2022 és nul.

²⁷⁴ Article 3.1.o de la Llei 15/2022.

²⁷⁵ Llei 19/2020, de 30 de desembre, d'igualtat de tracte i no-discriminació. Diari Oficial de la Generalitat de Catalunya, núm. 8307, de 31 de desembre de 2020. Aquesta llei tracta transversalment tot tipus de discriminacions (inclou a destacar com a característiques protegides –art. 1.3.h.- la condició social i econòmica, situació administrativa, professió o condició de privació de llibertat, i qualsevol manifestació d'aporofòbia o d'odi a les persones sense llar), deixant oberta la llista d'atributs concreta amb la següent fórmula: “*Qualsevol altra característica, circumstància o manifestació de la condició humana, real o atribuïda, que siguin reconegudes pels instruments de dret internacionals*” (art. 1.3.1). Es divideix en una part de foment que tracta les principals accions per evitar la discriminació en l'administració i la societat i una part sancionadora.

²⁷⁶ Article 19.1 del TFUE.

²⁷⁷ Article 19.2 del TFUE.

²⁷⁸ La Ley 15/2022, de 12 de julio, integral para la igualdad de trato y la no discriminación defineix en el seu article 6.3.a la discriminació múltiple com aquella que es produeix quan una persona es discriminada de manera simultània o consecutiva per dues o més causes de les previstes en la llei.

inusual que un algoritme discrimini únicament en base a un motiu protegit, ja que fonamenta els seus resultats en una multitud de factors diferents i variables estadísticament relacionades. Aquest fet qüestiona si la llista exhaustiva d'atributs protegits definits per l'article 19.1 del TFUE²⁷⁹ pot adaptar-se a les particularitats de la discriminació algorítmica.

Poden donar-se llavors situacions que no quedin cobertes per la normativa antidiscriminació, com que la categoria protegida no sigui suficientment representativa i causi exclusió de forma directa (1.2.1) o que es discrimini en base a motius no estipulats en la llista de categories protegides (1.2.2).

1.2.1. Atributs protegits no suficientment representatius

En referència a que una categoria sensible pugui no ser suficientment representativa²⁸⁰, s'ha de tenir en compte que les categories no es fan per representar la realitat sinó una manera de veure la realitat. S'explica bé aquest concepte en relació amb la discapacitat, ja que diversos sistemes de classificació (mèdic, científic, legal) s'interrelacionen amb institucions socials i les seves articulacions de poder i coneixement²⁸¹ per categoritzar persones.

S'ha destacat, per exemple, que la categorització basada en el sexe pot ocasionar l'exclusió de persones que s'identifiquen com a no binàries, transsexuals o intersexuals per accedir a certs béns, serveis o treballs o forçar-les a enquadrar-se en categories amb les quals no s'identifiquen, com també passa amb les tecnologies de control biomètric que forcen una perspectiva de binarisme de gènere²⁸².

²⁷⁹GERARDS, J. i XENIDIS, R., *Algorithmic Discrimination in Europe...*, op. cit., 2021, pàg. 9.

²⁸⁰SCHEUERMAN, M.K., WADE, K., LUSTING, C. i BRUBAKER, J.R., "How We've Taught Algorithms to See Identity: Constructing Race and Gender in Image Databases for Facial Analysis", *Proceedings of the ACM on Human-Computer Interaction*, vol. 4, article 58, 2020, pàg. 18. De les bases de dades analitzades en aquest article, gairebé totes només empraven les categories equivalents a "home" o "dona" per classificar el gènere. Només en una hi havia en la descripció de les anotacions de gènere un valor atípic: "5% roman insegur".

²⁸¹WHITTAKER, M. et al., *Disability, Bias and AI*, AI Now Institute, New York, November 2019; CRAWFORD, K., *Atlas of AI*, op.cit., 2021, pàg. 148. L'autora especifica que han esdevingut estàndard tant les classificacions no consensuals com les assumpcions normatives sobre la identitat.

²⁸²GERARDS, J. i XENIDIS, R., *Algorithmic Discrimination in Europe...*, op. cit., 2021, pàg. 62.

Aquí cal puntualitzar que en aquest treball, seguim la terminologia “discriminació per raó de sexe” i “igualtat salarial entre sexes” (i no gèneres) atès que és la seguida pel TJUE. El sexe seria el conjunt de modalitats bioquímiques, fisiològiques i orgàniques que polaritzen els individus d'una mateixa espècie en mascles i femelles (...) ²⁸³, i en canvi, el gènere serien el conjunt de característiques d'origen cultural relatives a patrons de comportament i d'identitat sobre la base de les quals s'estableix socialment la distinció entre homes i dones ²⁸⁴. Històricament s'ha lligat el sexe de les persones al gènere, no obstant això, aquesta relació actualment es posa en dubte. És una discussió rellevant: Arcos Vargas afirma que dins la discriminació per raó de sexe tal com entesa està pel TJUE s'engloba aquella discriminació en base a la “identitat de gènere”, apuntant seguidament, nogensmenys, que sembla que la legislació en relació a la “identitat de gènere” requereix considerable aclariment a escala tant europea com nacional ²⁸⁵.

En aquest sentit, s'ha puntualitzat que és incert com la legislació d'igualtat entre sexes de la UE protegeix de la discriminació les persones intersexuals o les persones que no s'identifiquen amb les normes de gènere que s'esperen d'elles, en no incloure conceptes com la identitat de gènere o l'expressió de gènere o trobar-se que el TJUE ha reforçat el binarisme de gènere en relació amb els drets de les persones transexuals en comptes de fer espai per identitats no binàries ²⁸⁶. Addicionalment, defensen que atès que els algoritmes evolucionen en el temps (com també ho fa la seva funció de reconeixement de patrons), depenent de l'algoritme –i agafant com a exemple el gènere–, ser membre d'una o més construccions de gènere pot ser dinàmic, ja que algoritmes diferents poden identificar i subratllar diferents característiques i preferències individuals ²⁸⁷.

1.2.2. Absència de protecció d'atributs que reforcen les desigualtats socials

²⁸³ Sexe. Gran enciclopèdia catalana. 2023. Disponible a: https://www.enciclopedia.cat/cerca/gec?search_api_fulltext=sexe&field_faceta_cerca_1=All. Darrera visita el 2 de setembre de 2023.

²⁸⁴ Gènere. TERMCAT Centre de Terminologia. 2023. Disponible a: <https://www.termcat.cat/es/cercaterm/fitxa/MjcwMTQ1NA%3D%3D>. Darrera visita el 2 de setembre de 2023.

²⁸⁵ ARCOS VARGAS, M., “La no discriminación en el derecho derivado...”, *op. cit.*, 2022, pàg. 27 i 28.

²⁸⁶ Vegeu: VAN DEN BRINK, M. i DUNNE, P., *Trans and intersex equality rights in Europe - a comparative analysis*, European network of legal experts in gender equality and non-discrimination, Brussels, 2018.

²⁸⁷ GERARDS, J. i XENIDIS, R., *Algorithmic Discrimination in Europe...*, *op. cit.*, 2021, pàg. 66.

En referència a la discriminació en base a motius no estipulats en la llista d'atributs protegits, doctrinalment s'ha volgut demostrar que la presa de decisions mitjançant algoritmes conté patrons de discriminació estructural que es basen en desigualtats relacionades amb característiques com l'estatus econòmic o els ingressos, i aquestes característiques no estan protegides pel dret derivat de la UE²⁸⁸.

La discriminació algorítmica té la capacitat d'estendre's a persones que no pertanyen a col·lectius protegits tradicionalment²⁸⁹. Com hem explicat en l'apartat 2.1.4 del capítol I en relació als indicadors, si el sistema d'intel·ligència artificial crea noves categories que no es relacionen amb les característiques protegides concretes (p. e. estatus socioeconòmic) però que igualment diferencien a les persones, aquestes categories poden ser injustes si reforcen les desigualtats socials²⁹⁰. Per exemple, l'origen social, el naixement i la propietat són tres característiques que poden estar relacionades amb la raça i l'ètnia i suposen un estatus imposat a la persona en virtut d'unes condicions socials, econòmiques i biològiques heretades, però cap d'elles es recull en les directives de la UE²⁹¹. Sempre hi ha debat sobre quins atributs han de quedar dins o fora d'aquells que necessiten especial protecció. Per exemple, actualment s'està debatent, si es podria incloure en els motius de discriminació els grups amb baix poder adquisitiu, baix nivell educatiu, tipus de vida no saludable o amb obesitat²⁹².

Així i tot, malgrat que es critica la concreció de categories protegides, com expliquen Barrère i Morondo, avui en dia no hi ha una altra manera instrumentalitzar el dret de manera antidiscriminatòria que fent ús de classificacions raonables²⁹³ (fet que suposa indefectiblement una exclusió de certes categories que no estiguin comunament reconegudes). A més (i com veurem més extensament en l'apartat 2.2 del capítol IV), l'existència de categories protegides té una incidència directa en la detecció de la

²⁸⁸ XENIDIS, R., "Turning EU equality law to algorithmic discrimination: Three pathways to resilience", *Maastricht Journal of European and Comparative Law*, vol. 27, núm. 6, 2020, pàg. 752.

²⁸⁹ *Ídem*, pàg. 747.

²⁹⁰ ZUIDERVEEN, F., *Discrimination, artificial intelligence...*, *op. cit.*, 2018, pàg. 7 i 8; MANN, M. i MATZNER, T., "Challenging algorithmic profiling: The limits of data protection and anti-discrimination in responding to emergent discrimination", *Big Data and Society*, vol. 6, núm. 2, 2019, pàg. 5.

²⁹¹ ARCOS VARGAS, M., "La no discriminación en el derecho derivado...", *op. cit.*, 2022, pàg. 31.

²⁹² GERARDS, J. i ZUIDERVEEN, F., "Protected Grounds and the System...", *op. cit.*, 2022, pàg. 44 i 45.

²⁹³ BARRÈRE, M.A. i MORONDO, D., "Subordinación y discriminación interseccional: elementos para una teoría del derecho antidiscriminatorio", *Anales de la Cátedra Francisco Suárez*, núm. 45, 2011, pàg. 28.

discriminació múltiple²⁹⁴, bàsicament perquè si la normativa no cobreix tots els motius de discriminació segons els quals la presumpta discriminació s'ha produït, només els motius coberts per la normativa poden ser tinguts en compte²⁹⁵.

En definitiva, la forma de discriminar mitjançant sistemes algorítmics fa palesa la limitació que suposa comptar amb una llista exhaustiva de motius a protegir per part de la legislació antidiscriminatòria, que per la seva configuració no només causa exclusió dins de cada categoria (p. e. persones no binàries, transsexuals o intersexuals) sinó que també exclou directament de la protecció moltes tipologies de diferenciació injustes que reforcen desigualtats socials (p. e. socioeconòmiques). Tanmateix, s'ha de tenir en compte que és a través de l'establiment de categories protegides que es pot instrumentalitzar el dret antidiscriminatori. Seguidament analitzem les limitacions que també s'han trobat en les directives antidiscriminació per abordar certs aspectes de la discriminació algorítmica.

1.3. Directives antidiscriminació europees

És a partir de l'article 13 del Tractat d'Amsterdam (actual article 19 TFUE) que es van promulgar les directives antidiscriminació europees, caracteritzades per la seva llista de motius exhaustiva, segons la qual la creació d'un desavantatge desproporcionat a una persona o grup identificat en base a un dels motius suposa discriminació. Aquests motius són el sexe, l'origen ètnic o raça, la religió o creença, la discapacitat, l'orientació sexual o l'edat.

La discriminació en relació a l'origen ètnic o racial està prohibida per la Directiva 2000/43/CE en assumptes de treball, ocupació i formació professional, i esferes no laborals com el benestar social, la protecció social, la seguretat social i l'assistència

²⁹⁴ *Ídem*, pàg. 23 i 40. Les autores expressen que malgrat que l'opressió tingui diverses cares (maneres de discriminar segons la categoria protegida), i hi hagi interseccions entre elles que ocasionin que els grups oprimits no resultin homogenis ni estancs, això no significa que es minimitzi la necessitat de comptar amb categories per identificar-la (classe, sexe-gènere, raça, etc.), o que aquestes es fragmentin donant lloc a infinits grups oprimits que perdin la seva identitat política grupal per dissoldre's en la individualitat. És a dir, si bé les categories d'anàlisi no tenen perquè funcionar com compartiments estancs, si que resulten imprescindibles.

²⁹⁵ TOBLER, C., *Indirect discrimination under Directives 2000/43 and 2000/78*, op. cit., 2022, pàg. 94-95.

sanitària, els beneficis socials, l'educació i l'accés a béns i serveis, inclòs l'habitatge²⁹⁶. L'ampliació a l'accés al sistema de benestar social i la seguretat social, així com als béns i serveis respon a la convicció que per tal que les persones puguin desenvolupar tot el seu potencial en el mercat laboral, resulta també essencial garantir la igualtat d'accés en àrees com la salut, l'educació i l'habitatge²⁹⁷. Estén el seu abast, a través de la Decisión marco 2008/913/JAI del Consejo, de 28 de noviembre de 2008, relativa a la lucha contra determinadas formas y manifestaciones de racismo y xenofobia mediante el Derecho penal²⁹⁸ que cobreix algunes de les formes més perjudicials de discriminació racial i xenòfoba, especificades en la incitació a la violència contra grups definits per la seva raça, color, religió, ascendència u origen nacional o ètnic i en la condonació, negació o banalització de delictes com el genocidi.

La discriminació per raó de sexe, té un ampli abordatge en el dret primari²⁹⁹, desenvolupant-se en la legislació derivada, entre d'altres³⁰⁰, en les Directives 2004/113/CE i 2006/54/CE. La Directiva 2004/113/CE fa referència a la implementació

²⁹⁶ Article 3 de la Directiva 2000/43/CE.

²⁹⁷ ARCOS VARGAS, M., "La no discriminación en el derecho derivado...", *op. cit.*, 2022, pàg. 21.

²⁹⁸ Decisión marco 2008/913/JAI del Consejo, de 28 de noviembre de 2008, relativa a la lucha contra determinadas formas y manifestaciones de racismo y xenofobia mediante el Derecho penal. DOUE, L 328/55, 6.12.2008.

²⁹⁹ L'article 2 del TUE estableix la igualtat entre dones i homes entre els valors sobre els quals es fonamenta la UE, i l'article 3.3 del TUE posiciona la igualtat de sexes com una de les missions de la UE (com també, combatre l'exclusió social i la discriminació). L'article 8 del TFEU integra la promoció de la igualtat entre dones i homes en totes les activitats de la UE, i es recolza en l'article 10 del TFUE que proclama la lluita envers la discriminació basada en el sexe (com també en l'origen racial o ètnic, la religió o la creença, la diversitat, l'edat o l'orientació sexual). L'article 19 del TFUE inclou el sexe com a un motiu protegit de discriminació sota la legislació europea. L'article 153.1.i del TFUE clarifica que en relació a polítiques socials, la UE hauria de donar suport i complementar les activitats dels estats membres en relació a la igualtat entre dones i homes d'oportunitats en el mercat laboral i de tracte a la feina. L'article 157 del TFUE conté provisions relatives a la igualtat en el salari d'homes i dones, igualtat de tractament a la feina i accions positives. A més a més, l'article 20 de la CDFUE proclama el principi d'igualtat davant la llei (que podríem dir que inclou la igualtat de sexes), l'article 21 i 23 del mateix text, prohibeix el sexe com a motiu de discriminació entre d'altres i proclama el principi d'igualtat entre homes i dones en totes les àrees, respectivament.

³⁰⁰ Directiva 79/7/CEE del Consejo, de 19 de diciembre de 1978, relativa a la aplicación progresiva del principio de igualdad de trato entre hombres y mujeres en materia de seguridad social. DOCE, L 6/24, 10.01.1979; Directiva 92/85/CEE del Consejo, de 19 de octubre de 1992, relativa a la aplicación de medidas para promover la mejora de la seguridad y de la salud en el trabajo de la trabajadora embarazada, que haya dado a luz o en período de lactancia. DOCE, L 348/1, 28.11.1992; Directiva 97/81/CE del Consejo, de 15 de diciembre de 1997, relativa al Acuerdo marco sobre el trabajo a tiempo parcial concluido por la Unice, el Ceep y la Ces. DOCE, L 14/9, 20.01.1998; Directiva 2010/18/UE del Consejo, de 8 de marzo de 2010, por la que se aplica el Acuerdo marco revisado sobre el permiso parental, celebrado por BUSINESSSEUROPE, la UEAPME, el CEEP y la CES. DOUE, L 68/13, 18.03.2010; Directiva 2010/41/UE del Parlamento Europeo y del Consejo, de 7 de julio de 2010, sobre la aplicación del principio de igualdad de trato entre hombres y mujeres que ejercen una actividad autónoma. DOUE, L 180/1, 15.07.2010; Directiva (UE) 2019/1158 del Parlamento Europeo y del Consejo, de 20 de junio de 2019, relativa a la conciliación de la vida familiar y la vida profesional de los progenitores y los cuidadores. DOUE, L 188/79, 12.07.2019.

del principi d'igualtat de tractament entre dones i homes en l'accés i subministrament de béns i serveis³⁰¹ de manera que tracta la igualtat de sexes en el mercat de consum, i cobreix qualsevol context en el qual es presti un bé o un servei normalment a canvi d'una remuneració, sempre que no es produeixi en un context enterament personal³⁰² (tal com la Directiva 2000/43/CE, que també cobreix l'àrea de béns i serveis).

Cal destacar que exclou de la seva aplicació l'àmbit dels mitjans de comunicació, la publicitat i l'educació³⁰³. Aquests buits suposen que la Directiva 2004/113/CE no pugui prevenir totalment la publicitat estereotipada, ofensiva i sexista³⁰⁴.

La Directiva refosa 2006/54/CE fa referència a la implementació del principi d'igualtat d'oportunitats i d'igualtat de tractament entre dones i homes en àmbits del treball i d'ocupació. Cal puntualitzar que només garanteix la igualtat de tracte en referència a la seguretat social i no al sistema de benestar social que inclou, per exemple, la protecció social, l'accés a la sanitat i a l'educació -prevista en la Directiva 2000/43/CE³⁰⁵.

L'article 3.2, tant de la Directiva 2000/43/CE com de la Directiva 2006/54/CE, estableix que no es cobreixen diferències de tracte basades en la nacionalitat³⁰⁶.

La discriminació per raó de religió o creença, discapacitat, edat i orientació sexual es protegeix mitjançant la Directiva 2000/78/CE que només s'aplica en assumptes laborals. Com a resultat, la discriminació per aquests motius no està expressament prohibida en l'àmbit de l'educació, la seguretat social, l'accés a béns i serveis incloent salut, habitatge, publicitat i mitjans de comunicació. S'estableix un marc general d'igualtat de tracte en el treball i l'ocupació concretant certes excepcions al principi d'igualtat. L'article 2.5 estableix que la Directiva s'entendrà sense perjudici a les mesures de la legislació nacional necessàries per a la seguretat pública, la defensa de l'ordre i la prevenció d'infraccions penals, la protecció de la salut i la protecció de drets i llibertats de la

³⁰¹ L'article 3.1 de la Directiva 2004/113/CE, precisa que fa referència a béns i serveis disponibles per al públic i que s'ofereixin fora de l'àmbit de la vida privada i familiar, i les transaccions que s'efectuïn en aquest context. La Directiva fa referència a l'article 50 del Tractat Constitutiu de la Comunitat Europea (actual article 57 TFUE), en el qual es consideren serveis als efectes del Tractat les prestacions realitzades a canvi de remuneració i comprendran en particular: a) activitats de caràcter industrial; b) activitats de caràcter mercantil; c) activitats artesanals; i d) activitats pròpies de les professions liberals.

³⁰² ARCOS VARGAS, M., "La no discriminación en el derecho derivado...", *op. cit.*, 2022, pàg. 36.

³⁰³ Article 3 de la Directiva 2004/113/CE.

³⁰⁴ WACHTER, S., "Affinity profiling and discrimination...", *op. cit.*, 2020, pàg. 392.

³⁰⁵ AGENCIA EUROPEA DE DERECHOS FUNDAMENTALES (FRA) y CONSEJO DE EUROPA, *Manual de legislación Europea contra la discriminación*, Luxemburgo, 2019, pàg. 23.

³⁰⁶ Article 3.2 de la Directiva 2000/43/CE i de la Directiva 2000/78/CE.

ciutadania³⁰⁷. Podrà no aplicar-se la directiva a les forces armades en allò que respecta a la discriminació basada en la discapacitat i l'edat³⁰⁸. Permet una diferència de tracte, quan els atributs protegits tenen relació intrínseca amb la capacitat de desenvolupar una feina determinada o les qualificacions determinades per realitzar-la³⁰⁹, o basada en la religió o conviccions d'una persona, en el cas d'activitats professionals d'esglésies i altres organitzacions públiques i privades l'ètica de les quals es basi en la religió o les conviccions, sempre i quan la diferència constitueixi un requisit professional essencial, legítim i justificat respecte l'ètica de l'organització, i la política s'ajusti als límits constitucionals dels estats membres i als principis generals del dret comunitari³¹⁰. Finalment, també permet una diferència de tracte basada en l'edat, per exemple, l'establiment d'una edat mínima per accedir a una feina sempre que estigui justificada per una finalitat legítima³¹¹.

Un dels problemes principals que podem detectar en la perspectiva que s'adopta per abordar la discriminació, és que, com hem vist, la protecció de la discriminació no és igual per a tots els atributs protegits. Només el sexe i l'origen ètnic, són atributs protegits més enllà del mercat (per exemple, seguretat social i sanitat en relació al sexe, i educació en relació a l'origen ètnic). En el mercat estan també protegits més àmpliament que els altres atributs (accés als béns i serveis en comptes de només contextos laborals). Això suposa que l'origen ètnic i el sexe estan protegits horitzontalment, és a dir, no només en les relacions publicoprivades que involucren l'Estat, sinó que també en la relació entre actors privats en el domini públic (activitats comercials i professionals), i en canvi, els altres atributs (la discapacitat, edat, orientació sexual, religió o creença) només es protegeixen en el context laboral³¹².

La manca de protecció enfront la discriminació basada en edat, discapacitat, orientació sexual i religió o creences en l'àrea de béns i serveis per part de la legislació antidiscriminatòria europea representa un buit en un moment en què els algorismes s'estan

³⁰⁷ Article 2.5 de la Directiva 2000/78/CE.

³⁰⁸ Article 3.4 de la Directiva 2000/78/CE.

³⁰⁹ Article 4.1 de la Directiva 2000/78/CE. També, de la mateixa manera, trobem aquesta exclusió en l'article 14.2 de la Directiva 2006/54/CE i l'article 4 de la Directiva 2000/43/CE.

³¹⁰ Article 4.2 de la Directiva 2000/78/CE.

³¹¹ Article 6 de la Directiva 2000/78/CE.

³¹² XENIDIS, R., *Shaking the normative foundations of eu equality law: evolution and hierarchy between market integration and human rights rationales*, Working Paper, EUI LAW, 2017/04, European Regulatory Private Law Project (ERPL), San Domenico di Fiesole, 2017, pàg. 25.

emprant per fer perfils de persones amb finalitats de venda³¹³, multiplicant les oportunitats que les companyies privades facin perfils de clientela potencial mitjançant ofertes personalitzades. La discriminació algorítmica és llavors molt més procliu a ocórrer en el mercat de béns i serveis, ja que el seu funcionament és analitzar el comportament de les persones usuàries per donar una resposta a cada una amb diferents ofertes i oportunitats individualitzades, de manera que duent-se a terme la discriminació, grups i persones pertanyents a col·lectius desavantatjats poden ser exclosos legalment de l'accés a certs béns i serveis, tan importants com l'habitatge, la salut o l'educació³¹⁴. Malgrat es podria legislar a escala nacional (ja que es poden promulgar règims més restrictius que els de la Directiva -que estableix un mínim obligatori-), no hi ha cap prohibició harmonitzada a escala europea³¹⁵. Quan els estats membres afegeixen elements que no estan establerts en les directives (per exemple, afegeixen categories protegides o àmbits de protecció), actuen fora de l'àmbit d'aplicació del dret de la UE³¹⁶.

A més, malgrat que la Directiva 2004/113/EC proporciona protecció en relació al sexe, com ja hem dit, els mitjans de comunicació, la publicitat i l'educació estan exclosos de la seva aplicació. En vista de l'increment d'ús d'intel·ligència artificial en aquests camps, aquestes excepcions poden portar a debilitats en relació a la capacitat de la legislació europea per fer front a la discriminació algorítmica³¹⁷. En el cas de la Directiva 2000/43/CE, no hi ha una excepció explícita però tampoc hi ha cap menció a la seva aplicació en els camps dels mitjans de comunicació i la publicitat (l'educació, en canvi, sí que està inclosa³¹⁸). Això suposa un problema, ja que es genera una disparitat de protecció segons el motiu causant de discriminació en la legislació europea³¹⁹.

En tractar-se de directives (actes legislatius en el quals s'estableixen objectius que tots els estats membres han de complir però que correspon a cada país elaborar les seves pròpies lleis sobre com assolir els objectius) cal apuntar que atès que deixen un ampli marge als estats membres per al seu desenvolupament³²⁰ es genera molta disparitat de

³¹³GERARDS, J. i XENIDIS, R., *Algorithmic Discrimination in Europe...*, *op. cit.*, 2021, pàg. 9.

³¹⁴XENIDIS, R., "Turning EU equality law to algorithmic discrimination...", *op. cit.*, 2020, pàg. 741.

³¹⁵GERARDS, J. i XENIDIS, R., *Algorithmic Discrimination in Europe...*, *op. cit.*, 2021, pàg. 58.

³¹⁶TOBLER, C., *Indirect discrimination under Directives 2000/43 and 2000/78*, *op. cit.*, 2022, pàg. 72.

³¹⁷GERARDS, J. i XENIDIS, R., *Algorithmic Discrimination in Europe...*, *op. cit.*, 2021, pàg. 58.

³¹⁸Article 3.1.g de la Directiva 2000/43/CE.

³¹⁹GERARDS, J. i XENIDIS, R., *Algorithmic Discrimination in Europe...*, *op. cit.*, 2021, pàg. 55.

³²⁰ARCOS VARGAS, M., "La no discriminación en el derecho derivado...", *op. cit.*, 2022, pàg. 39; CHOPIN, I. i GERMAINE, C., *A comparative analysis of non-discrimination law in Europe 2022*, European network of legal experts in gender equality and non-discrimination, Brussels, 2022, pàg. 67. Les

legislacions³²¹. La disparitat de legislacions i perspectives per abordar la regulació de la lluita contra la discriminació combinada amb el fet que els sistemes algorítmics no tenen límits territorials, és a dir, són essencialment globals, clarament disminueix la capacitat protectora de la normativa en l'àmbit de la UE.

En definitiva, els límits en l'aplicació de les directives, ja sigui perquè es refereixen a atributs protegits determinats, i no a totes les formes de discriminació (p. e. cap dels instruments de dret derivat de la UE citats inclou l'origen social, la propietat o qualsevol altre element que pugui apuntar a la condició socioeconòmica com un dels motius especialment protegits³²²), ja sigui perquè la protecció és diferent segons l'atribut protegit, ja sigui perquè es veuen limitades pel sector en el qual s'apliquen (majoritàriament en relacions laborals i no en altres àmbits)³²³, ja sigui perquè proporcionen excepcions que suposen un gran buit en relació a les decisions preses mitjançant algoritmes (com l'àmbit de la publicitat), són límits que ja suposaven llacunes en l'efectivitat de la normativa antidiscriminatòria però que arran de la introducció dels sistemes algorítmics esdevenen més rellevants³²⁴. La normativa europea també presenta una configuració molt orientada a entendre la discriminació com un acte puntual, aïllat, i no com un acte que respon a la presència en la societat de sistemes d'opressió i dominació. Que no s'estableixi una perspectiva que integri els sistemes d'opressió i dominació també suposa un punt feble en la normativa antidiscriminació europea per abordar la discriminació algorítmica, com veurem tot seguit.

autores estableixen que alguns estats membres de la UE proveeixen la mateixa protecció que s'estableix en les directives antidiscriminació, per a altres característiques, per tant, van més enllà dels requeriments de les directives (per exemple, França o Finlàndia).

³²¹ Per saber en cada estat membre de la UE quines característiques estan protegides per la legislació, vegeu: CHOPIN, I. i GERMAINE, C., *A comparative analysis of...*, *op. cit.*, 2022, pàg. 11-15.

³²² SORIANO ARNANZ, A., "Discriminación algorítmica: garantías y protección jurídica", a COTINO HUESO, L., (dir.) i BAUZA REILLY, M., (coord.), *Derechos y garantías ante la inteligencia artificial y las decisiones automatizadas*, Navarra, Aranzadi, 2022, pàg. 143.

³²³ *Ídem*, pàg. 163. L'autora especifica que moltes aplicacions dels sistemes algorítmics no entren en l'àmbit d'aplicació de les directives d'igualtat. Per exemple, l'ús d'algoritmes en motors de recerca que perpetuen certs estereotips que perjudiquen als membres de grups desavantatjats.

³²⁴ GERARDS, J. i XENIDIS, R., *Algorithmic Discrimination in Europe...*, *op. cit.*, 2021, pàg. 151-152. S'estableix en l'article que l'escala i la velocitat mai vista de la discriminació algorítmica pot ocasionar que s'agreugin les llacunes ja criticades anteriorment per la doctrina de la legislació antidiscriminatòria europea, a saber, els diferents àmbits d'aplicació material per les lleis antidiscriminatòries, la manca de reconeixement de la discriminació interseccional o la natura exhaustiva de la llista d'atributs protegits.

2. Discriminació algorítmica com a conseqüència de sistemes d'opressió sobre grups de persones

Per abordar la discriminació algorítmica com a conseqüència dels sistemes d'opressió estructurals cal relacionar, en primer lloc, la discriminació algorítmica amb les perspectives sobre la funció del dret antidiscriminatori, és a dir l'anticlassificació i l'antisubordinació (2.1), i en segon lloc evidenciar el caràcter grupal de la discriminació algorítmica, no abordat en un sistema de litigació marcadament individual (2.2).

2.1. Com posicionar la discriminació algorítmica dins la perspectiva del dret antidiscriminatori anticlassificació i antisubordinació

La funció del dret antidiscriminatori tradicionalment s'ha interpretat des de dues corrents diferents. D'una banda, la corrent anticlassificació defensa que el dret antidiscriminatori ha de posar l'atenció en què qualsevol tracte és igual quan és raonable i no arbitrari, rebutjant qualsevol tracte diferenciat entre persones sense guardar cap relació funcional amb la finalitat perseguida. Entén la discriminació com a trencament de la igualtat de tracte en termes bàsicament individuals, focalitzant-se en la igualtat formal i assumint una visió litigant i una concepció reparadora centrada en la dicotomia discriminació directa i indirecta (terme de comparació, justificació objectiva de la discriminació i litigació). En definitiva, suposa el tracte indiferenciat no discriminatori³²⁵.

Per exemple, des del punt de vista de la privacitat, l'anticlassificació prohibeix qualsevol diferenciació basada en un atribut protegit. El raonament darrera aquesta concepció és que si es pot prevenir l'accés a certa informació especialment protegida, es pot evitar la discriminació en base a aquesta informació protegida en un futur³²⁶. Vista des d'una

³²⁵ AÑÓN ROIG, M.J., "Desigualdades algorítmicas...", *op. cit.*, 2022, pàg. 41.

³²⁶ ROBERTS, J.L., "Protecting Privacy to prevent discrimination", *William and Mary Law Review*, vol. 56, núm. 6, 2015, pàg. 2123-2124.

posició crítica, aquesta perspectiva ignora la història de dominació i opressió que han patit certs grups de població generant un fals universalisme³²⁷.

D'altra banda, la corrent antisubordinació defensa que el dret antidiscriminatori té la finalitat de fer front a l'exclusió social, l'opressió i la subordinació de certs grups de població i s'ha d'articular al voltant de la noció d'igualtat entesa com a no sotmetiment o no exclusió. Aquesta perspectiva considera especialment rellevant l'existència de persones i grups de persones que, en haver estat històricament i sistemàticament excloses en diversos àmbits de la vida social, i haver-se vist subordinades i oprimides mitjançant estructures socials de poder i dominació, trenquen un esquema que es basa en una presumpta igualtat de totes les persones³²⁸. El trencament d'aquest esquema es produeix ja que, com a conseqüència d'aquesta situació de subordinació, les persones adscrites a aquests grups veuen limitades les seves oportunitats vitals i l'accés als béns i serveis produïts per la societat i es troben en una situació de desequilibri en el desenvolupament de les seves relacions socials³²⁹. L'antisubordinació es focalitza en la igualtat material, pretén eliminar el desavantatge patit pels grups oprimits a través d'estratègies de tracte diferencial per corregir la història de subordinació, estratègies de diversitat o polítiques que integrin les particularitats de cada grup en instruments reguladors generals³³⁰.

Aquesta concepció entén que el dret antidiscriminatori no pot prescindir de plantejaments que descansen sobre una determinada concepció del poder que s'ha denominat grupal, sistèmic i estructural i que respon a la idea d'opressió o "poder sobre", així com sobre persones excloses o subordinades, que adquireixen, en la seva consciència i en qualitat de tals, identitat política³³¹.

³²⁷BARRÈRE, M. A., "Filosofías del Derecho Positivo ¿Qué Derecho y qué discriminación? Una visión contrahegemónica del Derecho Antidiscriminatorio", *Anuario de Filosofía del Derecho*, vol. XXXIV, 2018, pàg. 13.

³²⁸ANÓN ROIG, M.J., "Desigualdades algorítmicas...", *op. cit.*, 2022, pàg. 41 i 42.

³²⁹ÁLVAREZ DEL CUVILLO, A., "El problema de la discriminación inversa: ¿Es posible discriminar a quienes pertenecen a los grupos sociales dominantes?", *Trabajo, Persona, Derecho, Mercado*, vol. 5, 2022, pàg. 189.

³³⁰SORIANO ARNANZ, A., *Data Protection for the prevention....*, *op. cit.*, 2021, pàg. 103; SORIANO ARNANZ, A., "Decisiones automatizadas y discriminación...", *op. cit.*, 2021, pàg. 10 i 11. L'autora considera que aquests instruments, com també els instruments que es concreten en les tècniques anticlassificatòries que es troben en l'article 9 de l'RGPD (prohibint el processament de categories de dades sensibles), suposen instruments de prevenció de la discriminació, i les prohibicions de discriminació directa i indirecta, com també mecanismes que pretenen evitar la presa de decisions discriminatòries, suposen instruments de reacció envers la discriminació.

³³¹BARRÈRE, M. A., "Filosofías del Derecho Positivo...", *op. cit.*, 2018, pàg. 20. L'autora també explica, en el mateix article, que l'enfocament d'aquesta perspectiva està orientat a identificar com es plasmen en

Els enfocaments de la discriminació en el CEDH i la jurisprudència del TEDH se situen clarament en l'òrbita del principi d'igualtat formal. El tractament de la discriminació per part de la normativa antidiscriminària de la UE -així com la jurisprudència del TJUE- és especialment ambigu i complex en la seva integració de les perspectives individual i col·lectiva: tot i que reconeix expressament conceptes com la discriminació indirecta i l'acció positiva que només poden entendre's des d'una perspectiva antisubordinatòria, al mateix temps tendeix a considerar les causes de discriminació com a classificacions contràries a un principi general d'igualtat de tracte entès en un pla abstracte i que per tant permet la tutela simètrica³³². De fet, la jurisprudència del TJUE admet obertament la possibilitat de discriminació inversa, atès que ha pres en consideració sovint les reclamacions basades en la suposada discriminació contra membres de grups dominants³³³.

La discriminació arran de l'anàlisi massiva de dades força una confrontació entre les dues perspectives que acabem d'esmentar. S'ha fet palès que la perspectiva anticlassificació pot no funcionar en els sistemes algorítmics. En l'anàlisi massiva de dades per no classificar persones en base a la seva pertinença a un o diversos grups, el model algorítmic hauria de, intencionadament, ignorar o suprimir tots els atributs protegits i tots els factors correlacionats amb els atributs protegits. Aquesta acció, ocasionaria un model amb resultats molt menys precisos³³⁴. A més, ens qüestionem si és possible desenvolupar aquest enfocament tenint en compte que en moltes ocasions no es poden determinar quines dades correlacionades acabaran proporcionant informació que la legislació considera protegida.

Barocas i Selbst plantegen, llavors, que donar solucions a la discriminació algorítmica per part de la legislació antidiscriminació requerirà un compromís amb els principis antisubordinació³³⁵, orientant-se la legislació cap a eliminar desigualtats basades en la

els sistemes jurídics aquelles desigualtats que tenen la seva raó de ser en les estructures de poder amb capacitat d'ordenar les relacions socials, per atribuir o assignar estatus (subordinats o inferiors, privilegiats o superiors) i per establir dinàmiques i inèrcies que reproduïxin aquestes relacions de subordinació (pàg. 32).

³³² ÁLVAREZ DEL CUVILLO, A., "El problema de la discriminación inversa...", *op. cit.*, 2022, pàg. 195.

³³³ *Ídem, ibidem*. Vegeu, per exemple: TJCE, sentència de 17 d'octubre de 1995, C-450/93, Eckhard Kalanke y Freie Hansestadt Bremen; TJCE (Sala Primera), sentència de 12 de desembre de 2019, C-450/18, WA contra Instituto Nacional de la Seguridad Social (INSS).

³³⁴ BAROCAS, S. i SELBST, A. D., "Big Data's Disparate...", *op. cit.*, 2016, pàg. 727.

³³⁵ AÑÓN ROIG, M.J., "Desigualdades algorítmicas...", *op. cit.*, 2022, pàg. 45. L'autora també apunta que el dret antidiscriminatori des de la concepció antisubordinació cobra especial rellevància en el

pertinença a certs grups socials, és a dir, en comptes de focalitzar-se en assegurar l'equitat processal, focalitzar-se en la substància, desenvolupant formes de raonament per jutjar quan i quina quantitat de discriminació és tolerable³³⁶. Aquesta visió està estretament lligada a la comprensió de la discriminació com a conseqüència de sistemes d'opressió que afecten grups de persones que comparteixen certa característica, no a una persona únicament. Com defensarem al llarg del treball, els mecanismes que es focalitzen en situacions concretes de discriminació perpetrades contra individus, fallen en assegurar la igualtat de tractament dels mateixos individus en risc de discriminació sistèmica i també fallen a protegir altres persones que poden patir danys col·laterals de les estructures socials que discriminen, de manera que estan mal equipats per fer front al dany massiu i les pràctiques sistèmiques il·legals³³⁷.

2.2. Discriminació algorítmica en base a grups estructuralment oprimits

La discriminació té una faceta o dimensió col·lectiva i grupal definitiva i no eliminable. La normativa antidiscriminària va sorgir de manera específica per erradicar determinades desigualtats sistèmiques que afectaven, i afecten, a determinats col·lectius perquè vulneraven, i vulnereu, la dignitat humana de les persones pertanyents a aquests col·lectius³³⁸.

El tracte desigual o injust és experimentat per persones individualment considerades, però la raó d'aquest tracte és que comparteixen o se'ls atribueixen unes característiques, trets

desenvolupament algorítmic, en la mesura que pot oferir respostes per abordar greus desigualtats estructurals.

³³⁶ BAROCAS, S. i SELBST, A. D., "Big Data's Disparate...", *op. cit.*, 2016, pàg. 723-728. S'alerta en l'article que l'espai entre la perspectiva antisubordinació i la legislació antidiscriminació cada cop és més gran, havent-se instaurat la perspectiva anticlassificació. En el context estatunidenc, els autors relaten que diversos litigis en referència a l'acció afirmativa han reescrit la distinció entre categoritzacions bones i perjudicials de raça a favor del principi formalista d'anticlassificació, suprimint la funció del dret antidiscriminació des dels seus orígens per protegir membres de grups històricament desavantatjats, fent que homes blancs actualment puguin denunciar que han patit discriminació (posen l'exemple del cas Ricci v. DeStefano).

³³⁷ FARKAS, L., "Collective actions under European anti-discrimination law", *European Anti-Discrimination Law Review*, vol. 19, 2014, pàg. 26.

³³⁸ ÁLVAREZ DEL CUVILLO, A., "El problema de la discriminación inversa...", *op. cit.*, 2022, pàg. 191. L'autor puntualitza que la prohibició de la discriminació que es basa en la dimensió col·lectiva sembla l'única perspectiva que es sosté des d'una interpretació històrica, sociològica, teleològica, sistemàtica i inclús literal del dret positiu vigent.

o prejudicis propis d'una col·lectivitat³³⁹ -independentment de la heterogeneïtat interna que es doni en el grup-, i que l'opressió envers el grup es tradueix en situacions de discriminació particulars³⁴⁰. La clàusula específica de no-discriminació en texts jurídics internacionals i constitucionals, pretén protegir determinats grups o col·lectius que, per la seva història de menysvaloració, subordinació i postergació social, no han format part, de fet, dels processos de formació de les normes³⁴¹.

La conducta discriminatòria es conforma per la confluència de dos elements constitutius: d'una banda, l'adscripció a un o diversos grups o categories socials, i d'altra banda, el perjudici individual i social provocat pel tractament derivat d'aquesta adscripció, que lesiona la dignitat humana en tant que tendeix a situar determinats grups en una posició sociojurídica d'inferioritat. Així doncs, només serien discriminatoris els tractaments desiguals, que estiguessin funcional o sistemàticament orientats a la marginació de determinats grups socials i les persones que els integren. Per tal que una conducta pogués considerar-se discriminatòria, s'exigiria que pogués contemplar-se com la manifestació d'una pauta sistemàtica de degradació de la posició dels integrants d'un grup que, si es generalitzés o es reproduís, contribuiria a situar-los en una posició d'inferioritat³⁴². Aquest seria un model teòric que partiria de la base que tot acte de discriminació ha d'implicar un perjudici social sobre determinats grups humans, assumint que aquest perjudici social pot ser només potencial³⁴³. En el moment en què s'admet que una norma pot tenir un transfons discriminatori, també s'admet que no es tracta només de protegir

³³⁹ AÑÓN ROIG, M. J., "Principio Antidiscriminatorio...", *op. cit.*, 2013, pàg. 134.

³⁴⁰ BARRÈRE, M.A., "Iusfeminismo y derecho antidiscriminatorio: hacia la igualdad por la discriminación", a MESTRE i MESTRE, R.M., (coord.), *Mujeres, derechos y ciudadanías*, Valencia, Tirant Lo Blanch, 2008, pàg. 60-62.

³⁴¹ AÑÓN ROIG, M. J., "Principio Antidiscriminatorio...", *op. cit.*, 2013, pàg. 131.

³⁴² ÁLVAREZ DEL CUVILLO, A., "El problema de la discriminación inversa...", *op. cit.*, 2022, pàg. 190, 191 i 196. L'autor exposa que la discriminació generalment admesa és un concepte híbrid que abasta tant una dimensió individual com una dimensió col·lectiva. No obstant això, malgrat que existeixen alguns punts de confluència entre ambdues tesis, en realitat són paradigmes contradictoris i mútuament excloents, fet que no impedeix que s'apliquin en la practica de manera combinada, entrant en nombroses incoherències i contradiccions lògiques. Una d'aquestes contradiccions és el concepte "discriminació inversa" que s'aplica als tractes que provoquen un perjudici a persones que pertanyen a grups socials dominants o majoritaris precisament per la seva adscripció a aquests col·lectius. Es tracta d'una apropiació per part dels grups dominants de la categoria jurídica de discriminació, que acaba per diluir el seu potencial transformador, legitimant així l'ordre de poder desigual. En aquest cas, no només la discriminació deixaria de complir la seva funció emancipadora, sinó que, a més, es convertiria en un instrument de dominació.

³⁴³ *Ídem*, pàg. 208.

persones discriminades sinó grups o col·lectius que pateixen un perjudici per les seves condicions personals o socials³⁴⁴.

Contràriament, la teorització del dret està presidida per la idea d'igualtat "de tots", volent donar el mateix estatus jurídic a les persones, sense reconèixer l'existència de grups subordinats. Aquesta igualtat política està basada en un fals universalisme -l'abstracció "individu" o "subjecte" en el dret està edificada i desenvolupada seguint un model concret (home, blanc, propietari...)-³⁴⁵. La idea que es promulga és que totes les persones tenen el mateix estatus jurídic (per tant, que el dret les considera iguals), i aquesta idea es manté sense fer diferències entre persones en la llei i en la seva aplicació, presumint-se una neutralitat per part de l'Estat. Quan les persones ja tenen els drets de manera igual, només es tracta d'intercedir perquè l'exercici abusiu (però individualitzat) d'uns no interfereixi en els drets (individuals) dels altres. En definitiva, la idea de justícia es tradueix en igualtat d'estatus jurídic individual (indiferenciació en la llei i indiferenciació en l'aplicació formal de la llei)³⁴⁶, com hem explicat en l'apartat 2.1 d'aquest capítol en base a la perspectiva anticlassificació.

Però el principi d'igualtat davant la llei (clàusula d'igual protecció) ignora determinats fenòmens estructurals d'injustícia social, plantejant un esquema en termes d'igualtat de tracte individual, no només sense resoldre la part estructural del conflicte sinó ajudant inclús a perpetuar la injustícia social³⁴⁷. La dissociació entre la normativa antidiscriminatòria i la lluita contra les desigualtats sistèmiques que realment operen en la societat, constitueix una distorsió manifesta de la seva finalitat que pot acabar desactivant la seva operativitat, atès que una visió purament individualista concep les causes de discriminació enunciades en les normes antidiscriminatòries simplement com classificacions abstractes³⁴⁸.

³⁴⁴ RODRÍGUEZ-PIÑERO y BRAVO-FERRER, M., "Los contornos de la discriminación", *Temas Laborales*, núm. 162, 2022, pàg. 15.

³⁴⁵ BARRÈRE, M.A., "Iusfeminismo y derecho antidiscriminatorio...", *op. cit.*, 2008, pàg. 54. Vegeu, en el mateix sentit: MACKINNON, C., "Difference and Dominance: on sex discrimination", a BARTLETT, K. T. i KENNEDY, R., (eds.), *Feminist legal theory: readings in law and gender*, Boulder, Westview Press, 1991, pàg. 82. L'autora exposa com davant l'estàndard en el qual tothom és igual, les dones es mesuren segons la correspondència amb els homes, de manera que el que se suposa neutralitat de gènere és simplement l'estàndard masculí.

³⁴⁶ BARRÈRE, M.A., "Iusfeminismo y derecho antidiscriminatorio...", *op. cit.*, 2008, pàg. 54-55.

³⁴⁷ *Ídem*, pàg. 58.

³⁴⁸ ÁLVAREZ DEL CUVILLO, A., "El problema de la discriminación inversa...", *op. cit.*, 2022, pàg. 192.

Aquest procés d'abstracció fa invisibles les desigualtats realment existents en la societat, projectant-les sobre el pla logicoformal desconnectat de la realitat material concreta. Aquesta forma de procedir permet concebre la discriminació com una situació anòmala, excepcional o patològica, ocultant d'aquesta manera situacions quotidianes que reproduïxen les relacions de dominació i subordinació, que acaba legitimant l'ordre de poder desigual³⁴⁹. En el mateix sentit apunten Barrère i Marondo, que la discriminació s'entén com un fenomen excepcional o accidental que requereix del dret una resposta, de la mateixa manera, puntual, reduint el concepte a les seves tipologies (directa i indirecta fonamentalment), quan no pot considerar-se com alguna cosa puntual o excepcional sinó com un fenomen sistèmic³⁵⁰, on la prohibició de discriminació opera com un mandat relativament autònom del principi d'igualtat, en el qual la parificació no es constitueix com una finalitat en si mateixa, sinó un mitjà al servei de l'eliminació de les desigualtats sistèmiques que vulneren la dignitat humana i que constitueixen un ordre polític i social manifestament injust³⁵¹. En aquest sentit és rellevant, com senyala Añón Roig, que els diferents enfocaments sobre el desenvolupament algorítmic (l'ètic, el tècnic i inclús el de risc) es caracteritzen per dues tendències, per una banda, trobar solucions circumscrites a la minimització de danys i per l'altra banda, assumir la premissa que els perjudicis que generen són individuals, quan en realitat també són socials i estructurals³⁵².

L'anàlisi massiva de dades, específicament l'elaboració de perfils, es basa en descobrir correlacions que es troben en bases de dades per identificar, representar i classificar a persones com a membres de grups o categories. A cada persona classificada en cada grup o categoria se li pot aplicar un perfil de grup (és a dir, característiques que l'algoritme, a partir de correlacions, considera probable que siguin comunes en les persones que integren el grup). Com més sofisticat esdevé el perfil de grup, donada la disponibilitat de -cada cop més- dades rellevants, més s'inclina cap a un perfil personalitzat i més

³⁴⁹ *Ídem*, pàg. 192 i 193.

³⁵⁰ BARRÈRE, M.A. i MORONDO, D., "Subordinación y discriminación interseccional...", *op. cit.*, 2011, pàg. 39. Especifiquen les autores que puntuals i excepcionals podran ser, en tot cas, les accions a les quals, segons el moment històric, se'ls reconeix o atribueix rellevància jurídica, però no els actes discriminatoris, que són, en definitiva, aquells que alimenten els sistemes d'opressió. Vegeu, en el mateix sentit: SEBASTIANI, L., "Investigando los límites de la lucha legal contra el racismo: El marco español de antidiscriminación por origen racial o étnico", *Oñati Socio-legal series*, vol. 11, núm. 3, 2021, pàg. 756. L'autor estableix que en relació amb el racisme, el debat hegemònic, malgrat que oscil·li entre considerar-lo com la conseqüència de perjudicis difosos o, alternativament, de predisposicions patològiques i extremistes, no deixa d'entendre'l com a acte "aïllat", desconnectat de dinàmiques socials més profundes, localitzable en la responsabilitat d'individus concrets.

³⁵¹ ÁLVAREZ DEL CUVILLO, A., "El problema de la discriminación inversa...", *op. cit.*, 2022, pàg. 192.

³⁵² AÑÓN ROIG, M.J., "Desigualdades algorítmicas...", *op. cit.*, 2022, pàg. 33.

subtilment discrimina entre membres i no membres del grup³⁵³. La creació d'aquests grups té la seva base en la reproducció d'estructures de privilegis socials i sistemes d'opressió.

Des de la perspectiva antilibertària, si l'anàlisi massiva de dades és el conjunt de tècniques i tecnologies que permeten explorar grans bases de dades, de manera automàtica i semiautomàtica, amb l'objectiu de trobar patrons repetitius, tendències o regles que expliquin el comportament de dades en un determinat context³⁵⁴, sempre que les dades estiguin lligades a comportament socials, els resultats que s'obtinguin a partir d'elles seran discriminatoris. Les tasques que duen a terme els sistemes algorítmics suposen l'aplicació d'aquests patrons repetitius, tendències o regles a conjunts de dades per extreure'n resultats. Aquests patrons, tendències o regles es donen en contextos socials determinats, de manera que el que fan els sistemes algorítmics bàsicament és reproduir sistemes de subordinació, opressió i exclusió. Per tant, en qualsevol ús de sistemes algorítmics predictius en el context social, s'obtidran resultats discriminatoris, atès que simplement els resultats reflectiran una societat desigual i discriminatoria i opressora de grups de població determinats que comparteixen certs atributs protegits.

En definitiva, la discriminació és la manifestació individualitzada de sistemes d'opressió i dominació -seguint la perspectiva antilibertària. L'opressió és una condició de grups i aquesta caracterització resulta fonamental en el concepte de discriminació³⁵⁵. Així i tot, actualment el dret antidiscriminatori de la UE està profundament marcat pel caràcter casuístic i individualitzat de l'aplicació judicial³⁵⁶ (que, entenem, facilita l'articulació de la protecció del dret en seu judicial, en construir el conflicte com un problema entre individus, però obviant que el conflicte és a causa de desigualtats estructurals). A més, la proposta de regulació anticlassificació que consisteix en no utilitzar dades sensibles en el sistema algorítmic pot no funcionar si aquestes s'infereixen amb altres dades. La discriminació causada per sistemes algorítmics fa palès el caràcter grupal de la discriminació en elaborar-se perfils en base a sistemes d'opressió i que aquests s'apliquin individualment a persones que pateixin una afectació per aquesta aplicació, és a dir, una

³⁵³ SCHREURS, W., "Cogitas, Ergo Sum. The role of data protection...", *op. cit.*, 2008, pàg. 242.

³⁵⁴ RIVAS VALLEJO, P., *La aplicación de la Inteligencia Artificial del trabajo y su impacto discriminatorio*, Navarra, Aranzadi, 2020, pàg. 103.

³⁵⁵ BARRÈRE, M.A. i MORONDO, D., "Subordinación y discriminación interseccional...", *op. cit.*, 2011, pàg. 20.

³⁵⁶ *Idem*, pàg. 37.

discriminació. El dret antidiscriminatori, com hem dit, potencia una lluita individual i fins que no es reconegui la necessitat de sistemes que apreciïn la incidència de l'opressió, sembla que no es podrà proporcionar una protecció adequada³⁵⁷.

3. Deficiències de la conceptualització bidimensional entre discriminació directa i discriminació indirecta³⁵⁸

El TEDH i el TJUE prohibeixen tant la discriminació directa com la discriminació indirecta³⁵⁹. En la discriminació directa l'element decisiu és la causa, i en canvi, en la discriminació indirecta l'element decisiu és l'efecte (essent necessari per part demandant plantejar un esquema comparatiu)³⁶⁰. Mentre la discriminació directa s'empra per establir si una persona ha estat objecte de discriminació, la discriminació indirecta es focalitza en la discriminació sistèmica de grup³⁶¹. La discriminació intencional usualment s'analitza emprant la doctrina de la discriminació directa, mentre que la discriminació no intencional però evitable s'aborda mitjançant la doctrina de la discriminació indirecta³⁶² (així i tot, ni en la discriminació directa ni en la discriminació indirecta s'exigeix provar una motivació de discriminar per apreciar discriminació). Distingir-les és important en relació a aspectes pràctics, per exemple, els motius per justificar la discriminació directa són *numerus clausus* i, en canvi, en la discriminació indirecta s'ha d'estudiar específicament cada cas³⁶³.

³⁵⁷ AÑÓN ROIG, M.J., “Desigualdades algorítmicas...”, *op. cit.*, 2022, pàg. 43 i 44. L'autora precisa que encara no s'ha produït un compromís clar del dret antidiscriminatori amb el principi antisubordinació i que és clarament dominant la perspectiva que concep el principi d'igualtat de tracte com a indiferenciació o tracte raonable no arbitrari (principi d'anticlassificació).

³⁵⁸ Als Estats Units, la discriminació indirecta s'anomena “disparate impact”, i en canvi, la discriminació directa s'anomena “disparate treatment”.

³⁵⁹ En el cas de discriminació directa, vegeu per exemple: TEDH (Secció Segona), sentència de 2 de desembre de 2014, 61960/08, Emel Boyraz contra Turquia; TJCE, sentència de 10 de juliol de 2008, C-54/07, Centrum voor gelijkheid van kansen en voor racismebestrijding contra Firma Feryn NV. En el cas de la discriminació indirecta, vegeu per exemple: TEDH (Secció Segona), sentència de 2 de febrer de 2016, 7186/09, Di Trizio contra Suïssa; TJCE, sentència de 13 de maig de 1986, C-170/84, Bilka - Kaufhaus GmbH contra Karin Weber von Hartz.

³⁶⁰ MALISZEWSKA-NIENARTOWICZ, J., “Direct and Indirect Discrimination in the European Union Law. How to draw a Dividing Line?”, *International Journal of Social Sciences*, vol. 3, núm. 1, 2014, pàg. 43.

³⁶¹ PÁEZ, A., “Negligent Algorithmic Discrimination”, *op. cit.*, 2021, pàg. 26.

³⁶² *Ídem*, pàg. 19.

³⁶³ TOBLER, C., *Indirect discrimination under Directives 2000/43 and 2000/78*, *op. cit.*, 2022, pàg. 78.

Cal puntualitzar que, des de la perspectiva antilibertària, es considera que la classificació directa/indirecta es va integrar en el dret europeu i va cristal·litzar en un esquema binari que obstaculitza l'avenç del dret antidiscriminatori i no facilita l'articulació de respostes enfront la discriminació algorítmica³⁶⁴. Es reclama que en el raonament sobre la no discriminació es donen elements que obliguen a superar la visió dicotòmica sobre les classes de discriminació, comprenent que la distinció entre discriminació directa i indirecta resulta limitada i insuficient per captar els patrons socials de discriminació o estructures de dominació³⁶⁵. S'alerta sobre la inconsistència del plantejament binari de la distinció tant des del punt de vista teòric (atès que tracta de conceptes heterogenis), com ideològic atès que eclipsa pel dret les relacions de poder intergrupals³⁶⁶. Persisteixen debats oberts sobre aquesta classificació, que Añón Roig considera exhaustiva i exclouent, però que és estàndard i està assentada en la legislació i en la jurisprudència³⁶⁷.

Leese planteja en relació a l'elaboració de perfils mitjançant algorismes, extrapolable a les decisions automatitzades, que la diferència entre aquests sistemes i la manera com s'havien dut a terme aquestes funcions tradicionalment, resideix en el fet que els mecanismes que no integraven algorismes entraven dins l'abast del marc regulador antidiscriminació per naturalesa, en canvi, les noves formes aconsegueixen escapar constantment dels règims reguladors vigents a causa de la fluïdesa dels algorismes adaptatius³⁶⁸. Les regulacions antidiscriminació només despleguen la plenitud del seu poder regulador quan s'apliquen a modes tradicionals de discriminació estàtics, i no a aquells provocats per l'anàlisi automatitzat de dades, dinàmic i implícit, que escapa constantment de la regulació i potser ja no es pot abordar mitjançant la perspectiva de la discriminació directa i indirecta establerta fins ara³⁶⁹.

³⁶⁴ AÑÓN ROIG, M.J., “Desigualdades algorítmicas...”, *op. cit.*, 2022, pàg. 36 i 37.

³⁶⁵ AÑÓN ROIG, M. J., “Principio Antidiscriminatorio...”, *op. cit.*, 2013, pàg. 130.

³⁶⁶ BARRÈRE, M.A., “Iusfeminismo y derecho antidiscriminatorio...”, *op. cit.*, 2008, pàg. 64. Sobre les relacions de poder, vegeu: FOUCAULT, M., *Microfísica del poder*, Madrid, Ediciones la Piqueta, 1979. L'autor exposa que no convé partir d'un fet primer i massiu de dominació (una estructura binària composta per “dominants” i “dominats”), sinó que cal considerar la producció multiforme de relacions de poder que són parcialment integrables a estratègies de conjunt. Les relacions de poder no donen suport a un interès econòmic primigeni sinó que són emprades i reajustades en el marc d'estratègies generals (pàg. 171).

³⁶⁷ AÑÓN ROIG, M.J., “Desigualdades algorítmicas...”, *op. cit.*, 2022, pàg. 36.

³⁶⁸ LEESE, M., “The New Profiling: Algorithms, Black Boxes...”, *op. cit.*, 2014, pàg. 498.

³⁶⁹ *Ídem*, pàg. 500.

En aquest apartat estudiem l'abordatge de la discriminació algorítmica, en primer lloc, a través de la doctrina de la discriminació directa (3.1), en segon lloc, a través de la discriminació indirecta (3.2), i en tercer lloc, analitzem per què sembla que les discriminacions algorítmiques hauran de ser desafiades majoritàriament com a discriminacions indirectes quan es poden donar tant directament com indirectament (3.3).

3.1. Discriminació algorítmica directa

La discriminació directa és aquella situació en la qual una persona és tractada de manera menys favorable del que ha sigut o vagi a ser tractada una altra persona en una situació comparable a causa d'una categoria protegida (per exemple l'ètnia o el sexe)³⁷⁰. Per determinar que existeix discriminació directa, el TEDH, requereix una diferència de tractament de persones en situacions anàlogues o significativament similars, basada en una característica identificable. Una definició similar s'empra per la legislació antidiscriminació de la Unió Europea³⁷¹.

Per justificar la discriminació directa, en el cas del TEDH només es requereix que es persegueixi una finalitat legítima i que els mitjans per aconseguir-la siguin adequats i necessaris. En el cas del dret de la UE, no hi ha una justificació objectiva general com en el TEDH, sinó que el tractament més desfavorable s'ha de justificar en base a les excepcions relatives a requisits professionals essencials³⁷². Aquestes excepcions permeten a aquelles persones que contracten diferenciar entre les persones pels seus atributs protegits (vegeu apartat 1.3 d'aquest mateix capítol).

El concepte de discriminació directa té certes limitacions que Fredman concreta en ser “relatiu” (requereix normalment un comparador apropiat) i “simètric” (també pot ser

³⁷⁰REY MARTÍNEZ, F., *Derecho Antidiscriminatorio*, op. cit., 2019, pàg. 56.

³⁷¹ZUIDERVEEN, F., “Strengthening legal protection against discrimination by algorithms and artificial intelligence”, *The International Journal of Human Rights*, vol. 24, núm. 10, 2020, pàg. 1576.

³⁷² AGENCIA EUROPEA DE DERECHOS FUNDAMENTALES (FRA) y CONSEJO DE EUROPA, *Manual de legislación Europea contra la discriminación*, Luxemburgo, 2019, pàg. 105 i 106.

invocat per membres de la majoria social en el cas de rebre un tracte diferent i pitjor que els d'una minoria social, només pel fet de pertànyer a la majoria)³⁷³.

Es tracta d'una forma de discriminació que es pot trobar quan es veuen implicats sistemes algorítmics, per exemple, quan aquests decideixen que les dones no són adequades per a un lloc de treball, ja que l'algoritme ha estat entrenat amb bases de dades on no hi ha pràcticament currículums de dones. Per tant, l'algoritme atorgarà al fet de ser dona una valoració menys positiva en comparació amb el fet de ser home, a causa de la seva infrarepresentació en la base de dades, i arribarà a la conclusió que es preferible contractar a homes. En conseqüència, el valor negatiu (o menys positiu) associat a la categoria de dades podrà determinar de manera directa el resultat³⁷⁴.

En la discriminació directa, el sistema algorítmic tracta una categoria protegida o la pertinença a un grup desavantatjat com a factor d'entrada negatiu per a la decisió a adoptar. També es pot donar quan el propi algoritme infereix la pertinença a un grup desavantatjat a través d'altres dades a les que atribueix un valor negatiu (ja sigui que les inferències les articulin les persones que creïn el programa, ja sigui que l'algoritme les desenvolupi mitjançant l'aprenentatge automàtic -discriminació a les dones a causa d'una base de dades integrada majoritàriament per homes-)³⁷⁵.

La discriminació directa, a més, es pot donar en dues tipologies. O bé la pertinença d'una persona al grup protegit és l'únic factor que es té en compte i que genera de manera automàtica un tracte discriminatori, o bé la pertinença d'una persona a un grup protegit no és l'únic factor que es té en compte i altres dades fan que el resultat final no sigui discriminatori. En aquest segon cas, hi ha una combinació de factors que fan que el resultat discriminatori relatiu a la categoria protegida es vegi compensat per altres factors. Per exemple, en el cas del sistema AMS, ja esmentat, ser dona resta punts, però ser

³⁷³ FREDMAN, S., *Discrimination Law*, Oxford, Oxford University Press, 2011, pàg. 166. L'autora es refereix quan parla de simetria a la discriminació inversa.

³⁷⁴ SORIANO ARNANZ, A., "Decisiones automatizadas y discriminación...", *op. cit.*, 2021, pàg. 15.

³⁷⁵ SORIANO ARNANZ, A., "Discriminación algorítmica: garantías y protección jurídica", *op. cit.*, 2022, pàg. 150. L'autora exposa que és rellevant si la dada és indicativa, però no determinant, de la pertinença al grup. En aquest cas és dubtós si es consideraria discriminació directa (factor equivalent a l'atribut protegit) o discriminació indirecta (mesura discriminatòria que només es pot provar a través dels resultats desavantatjats per un grup de població). Per exemple, haver anat a una escola en la qual la majoria de l'alumnat pertany a una minoria ètnica implica una major probabilitat de formar part d'aquell grup, però no és una variable totalment predictiva de la pertinença al grup si l'alumnat no es compon exclusivament de membres d'aquella ètnia minoritària. Una altra cosa seria la inferència de la pertinença al grup protegit a través de dades que siguin inseparables de la pertinença. Si resulta que a l'escola únicament hi accedeixen membres d'un determinat grup ètnic o racial, la dada serà inseparable de la pertinença al grup.

austríaca o nacional de la UE en suma³⁷⁶. Cal apuntar, com defensem al llarg d'aquest treball, que en el cas dels sistemes algorítmics on bàsicament hi ha anàlisi massiva de dades i correlació de dades és molt improbable que la decisió es basi únicament en una variable, com també serà difícil determinar quin pes es dona a les categories protegides i, si com a resultat del pes que se'ls dona, hi ha una afectació en el resultat final (on hi ha hagut discriminació però el resultat final pot acabar essent no discriminatori). Per tant, el problema principal de la vessant directa de la discriminació algorítmica serà provar que la tipologia de discriminació ha estat directa. Com estableix Benedi Lahuerta en relació amb el cas CHEZ, en els casos de discriminació algorítmica serà preferible demostrar l'existència de discriminació directa (a efectes de la restricció en la justificació posterior) però hi haurà d'haver suficients evidències per demostrar que la discriminació va produir-se *a causa de* la característica protegida³⁷⁷. Tot seguit abordem la discriminació algorítmica indirecta.

3.2. Discriminació algorítmica indirecta

La idea de discriminació indirecta està orientada a fer una valoració material de les desigualtats³⁷⁸. És una creació del Tribunal Suprem Federal dels Estats Units d'Amèrica del Nord en la sentència Griggs contra Duke Power Company, de 8 de març de 1971³⁷⁹. El debat residia en si era conforme a la constitució i a la US Civil Rights Act de 1964 que una empresa exigís per contractar i/o promocionar, haver acabat estudis secundaris i/o haver superat dos tests d'intel·ligència general, pràctica que perjudicava a la comunitat negra, ja que les estadístiques revelaven que si bé el 34% dels blancs completaven els estudis secundaris, només ho feia el 12% dels negres; i que si bé el 58% dels blancs superaven els tests d'intel·ligència utilitzats per l'empresa, només el 6% dels negres ho feia. El tribunal va observar que les mesures no tenien res a veure amb com es duia a terme la feina, raonant que amb caràcter general, una pràctica com aquesta, encara que

³⁷⁶ *Ídem*, pàg. 152.

³⁷⁷ BENEDI LAHUERTA, S., "Ethnic discrimination, discrimination by association and the Roma community: CHEZ", *Common Market Law Review*, vol. 53, núm. 3, 2016, pàg. 807 i 808.

³⁷⁸ AÑÓN ROIG, M. J., "Principio Antidiscriminatorio...", op. cit., 2013, pàg. 153.

³⁷⁹ Tribunal Suprem dels Estats Units d'Amèrica del Nord, sentència de 8 de març de 1971, Griggs v. Duke Power Co., 401 U.S. 424. Disponible a: <https://supreme.justia.com/cases/federal/us/401/424/>. Darrera visita el 2 de setembre de 2023.

fos neutral i no pretengués discriminar, no es podia conservar si servia per mantenir l'*statu quo* d'anteriors costums discriminatòries a la feina, donant rellevància a les conseqüències de la política empresarial i no només a la seva intenció³⁸⁰. En aquest sentit, Adams-Prassl, Binns i Kelly-Lyth afirmen que els circuits de retroalimentació³⁸¹ que es justifiquen a si mateixos bàsicament el que fan és mantenir un *statu quo* discriminatori³⁸², com per exemple en el cas de l'excés de vigilància policial en certs barris (vegeu apartat 2.1.1.2 del capítol I).

En conseqüència, la discriminació indirecta es podria definir com aquella situació en la que una pràctica que aparentment és neutra acaba discriminant persones dotades amb un atribut protegit³⁸³. El TJUE va començar a desenvolupar el concepte de discriminació indirecta en un cas d'igualtat de gènere a principis de la dècada de 1980, inspirat en la doctrina estatunidenca³⁸⁴. El TEDH, per la seva banda, la defineix com efectes desproporcionadament perjudicials d'una política o mesura general que, malgrat formulada en termes neutres, discrimina a un grup³⁸⁵. No es requereix necessàriament una voluntat discriminatòria de manera que no és rellevant si la persona presumptivament discriminadora té intenció de discriminar, sinó que és l'efecte de la pràctica el que compta, de manera que les persones que discriminen no poden evadir la prohibició de discriminació indirecta provant que no tenien la intenció de fer-ho³⁸⁶.

Un exemple a de discriminació indirecta a través de sistemes algorítmics seria el següent: la persona (o persones) a càrrec del sistema algorítmic, o el sistema algorítmic mitjançant l'aprenentatge automàtic, estableix que una de les característiques que cal observar per

³⁸⁰És aquest un clar exemple de tractar a persones en situacions diferents de manera idèntica, el que s'entén com a discriminació per indiferenciació o igualació.

³⁸¹ Els circuits de retroalimentació (en anglès feedback loop) es formen a través de perjudicis d'assignació i de representació causats pels biaixos i la discriminació existents, que poden tenir dinàmiques que es reforcen i assolixen velocitat desconeguda, opacitat i efectivitat en la presa de decisió mitjançant sistemes algorítmics.

³⁸² ADAMS-PRASSL, J., BINNS, R. i KELLY-LYTH, A., "Directly Discriminatory Algorithms", *Modern Law Review*, vol. 86, núm. 1, 2023, pàg. 154.

³⁸³Article 2.2.b. de la Directiva 2000/43/CE; article 2.2.b. de la Directiva 2000/78/CE; article 2.b. de la Directiva 2004/113/CE i; article 2.1.b. de la Directiva 2006/54/CE.

³⁸⁴ Vegeu: TJCE, sentència de 31 de març de 1981, C-96/80, JP Jenkins contra Kingsgate (Clothing Productions) Ltd; TJCE, sentència de 13 de maig de 1986, C-170/84, Bilka - Kaufhaus GmbH contra Karin Weber von Hartz.

³⁸⁵TEDH (Gran Sala), sentència de 24 de maig de 2016, 38590/10, Biao contra Dinamarca, par. 103. També, de manera similar en la regulació de la Unió Europea: article 2(2)(b) de la Directiva 2000/43/CE; article 2(2)(b) de la Directiva 2000/78/CE; article 2(1)(b) de la Directiva 2006/54/CE i; article 2(b) de la Directiva 2004/113/CE.

³⁸⁶TEDH (Gran Sala), sentència de 24 de maig de 2016, 38590/10, Biao contra Dinamarca, par. 103.

classificar com a “bona” a cada persona candidata a un lloc de treball és “no haver interromput la seva carrera professional en cap moment”. L’algoritme, en aquest cas, estaria causant discriminació indirecta contra les dones, atès que malgrat el fet de “no haver interromput la seva carrera professional en cap moment” sembla un criteri neutre, té un afectació més acusada en el cas de les dones que dels homes. Aquesta afectació es deu a que com a conseqüència de la maternitat i també de les tasques de cures que els són assignades en base a un sistema de divisió de rols per raó de sexe, les dones tendeixen a interrompre la seva carrera professional amb més freqüència que els homes³⁸⁷.

En base a la definició que s’estableix en les directives antidiscriminació³⁸⁸, podem determinar quatre requisits per trobar-nos davant d’una situació de discriminació indirecta (essent a partir d’aquests que analitzem les deficiències d’aquest concepte per abordar la discriminació algorítmica): un criteri, disposició o pràctica aparentment neutre (3.2.1), que creï un desavantatge particular d’un grup sobre un altre (3.2.2), que no sigui justificable objectivament amb una finalitat legítima (3.2.3) i que els mitjans per aconseguir-la siguin adequats i necessaris (3.2.4).

3.2.1. Criteri aparentment neutre

El fet de requerir que existeixi un criteri o pràctica aparentment neutra suposa que l’activitat ha d’estar habilitada per l’existència d’un tipus de requisit aplicable a totes o a un conjunt ampli de persones, que s’han de veure afectades pel seu impacte. Per poder obtenir protecció enfront la discriminació en base a la regulació europea és rellevant el factor relatiu a l’abast de la llei. El tribunal s’ha de qüestionar si el cas es pot integrar en el camp d’aplicació de la llei antidiscriminatòria europea que han d’aplicar els estats membres. El TJUE defensa que els elements que descriuen l’abast de manera positiva s’han d’interpretar àmpliament i, en canvi, els elements que descriuen l’abast de manera negativa s’han d’interpretar restringidament. Malgrat això, depèn molt de la manera com es posa en pràctica i la legislació dels estats membres a causa que aquests tenen la

³⁸⁷ SORIANO ARNANZ, A., “Discriminación algorítmica: garantías...”, *op. cit.*, 2022, pàg. 140.

³⁸⁸ Article 2.2.b. de la Directiva 2000/43/CE; article 2.2.b. de la Directiva 2000/78/CE; article 2.b. de la Directiva 2004/113/CE i; article 2.1.b. de la Directiva 2006/54/CE.

capacitat de regular la proporcionalitat entre la serietat de la intervenció i la gravetat de les raons que la justifiquen³⁸⁹. Per consegüent, s'ha de considerar que la manera com es dugui a terme la discriminació, -i la manera com la legislació estatal ha decidit abordar la qüestió, és a dir, que existeixi o no un criteri o pràctica aparentment neutra- determinaran si s'està dins l'abast de la protecció europea o no. Per exemple, casos que incloguin l'accés a un servei (assegurança) estaran coberts si s'al·lega discriminació en base al sexe o la raça, però no en relació a la discapacitat o a l'edat³⁹⁰.

És rellevant també si la presumpta víctima de discriminació indirecta pot provar que hi ha discriminació sota un motiu particular. Creem conflicte els casos en els quals la discriminació és a causa de diversos motius, ja que només aquells casos coberts per la llei és poden tenir en compte. Això comporta que es causi discriminació factual (factual discrimination), és a dir, pràctiques que, tot i no estar legalment prohibides, són discriminatòries en base a estàndards diferents, extralegals³⁹¹. Per tant, és pertinent remarcar que sense la concurrència d'un motiu protegit, el criteri o pràctica neutre suposarien establir desigualtat, no discriminació³⁹².

En les decisions automatitzades, la disposició, pràctica o criteri aparentment neutre poden ser, per exemple, els algoritmes usats per prendre la decisió o les pràctiques que es duen a terme sobre les dades que serviran d'entrenament per a l'algoritme³⁹³. Com la gran quantitat de dades tractades i la complexitat dels sistemes de tractament impedeixen identificar de manera clara quins són els elements que condicionen els resultats en el sistema, apunta Hacker que serà la totalitat del sistema el que haurà de ser considerat com una disposició, pràctica o criteri aparentment neutre que genera resultats discriminatoris³⁹⁴. D'aquesta manera, és més fàcil d'abordar la causa de la discriminació, en comptes de buscar dins el funcionament d'un sistema inherentment opac.

³⁸⁹TOBLER, C., *Limits and potential of the concept of indirect discrimination*, European Network of Legal Experts in Anti-Discrimination, Brussels, 2008, pàg. 38 i 68.

³⁹⁰*Idem*, pàg. 38.

³⁹¹*Idem, ibidem*.

³⁹²PUIG HERNÁNDEZ, M. A., *La protección de la libertad ideológica...*, *op. cit.*, 2018, pàg. 48.

³⁹³ ALLEN, R. i MASTERS, D., "Artificial Intelligence: the right to...", *op. cit.*, 2020, pàg. 592.

³⁹⁴HACKER, P., "Teaching fairness to artificial intelligence: existing and novel strategies against algorithmic discrimination under EU law", *Common Market Law Review*, vol. 55, núm. 4, 2018, pàg. 1161 i 1162.

3.2.2. Demostrar un desavantatge particular, sobretot en relació amb les proves estadístiques

Per determinar que ha existit discriminació indirecta s'ha de demostrar haver patit un desavantatge particular en relació a altres persones. En la discriminació directa, la persona demandant ha de provar que va ser tractada menys favorablement en base a una atribut protegit, provant un patró il·legal, com per exemple, que una companyia no contracta persones d'una ètnia particular, malgrat una part de la població sigui d'aquesta ètnia³⁹⁵. En canvi, la discriminació indirecta comporta un ús més extens de les estadístiques perquè els atributs protegits no són emprats explícitament³⁹⁶. Com la discriminació indirecta es centra en el resultat més que en el tracte, cal evidenciar l'impacte de la mesura que es posa en qüestió entre dos grups, el de la presumpta víctima i un altre. En els sistemes algorítmics, una vegada que l'algoritme incorpora un biaix que perjudica a les persones que pertanyen a un grup desavantajat, llevat que incorpori instruccions per corregir-lo - fet que succeeix de manera poc habitual-, es perpetua durant tot el període de funcionament del sistema³⁹⁷.

Suposant que fossin accessibles les dades, per demostrar un desavantatge particular, la prova estadística sembla l'única suficientment exhaustiva per abordar la discriminació arran de decisions automatitzades que operen analitzant dades massivament. Les evidències estadístiques poden ser indispensables per establir una possible correlació entre els atributs protegits legalment i els factors considerats en la norma impugnada. Mitjans tradicionals de prova, com la documental, testifical o opinions d'experts sovint no són efectives en els casos de discriminació indirecta³⁹⁸, especialment quan la relació entre un atribut protegit i un atribut no protegit s'ha de demostrar. Atès que la discriminació en base a sistemes automatitzats potser només és observable estadísticament, aquesta tipologia de prova pot ser la preferent en un futur.

³⁹⁵MAKKONEN, T., *Measuring Discrimination: Data Collection and EU Equality Law*, European Network of legal experts in the Non-discrimination field, Brussels, 2007, pàg. 32.

³⁹⁶*Idem*, pàg. 31.

³⁹⁷SORIANO ARNANZ, A., "Decisiones automatizadas y discriminación...", *op. cit.*, 2021, pàg. 21.

³⁹⁸FARKAS, L. i O'FARREL, O., *Reversing the burden of proof: practical dilemmas at the European and national level*, European Network of legal experts in the Non-discrimination field, Brussels, 2015, pàg. 24.

S'ha recorregut a la prova estadística tant per part de TJUE com per part de TEDH. En l'assumpte Seymour-Smith i Pérez, el TJUE especifica que correspon a l'òrgan jurisdiccional nacional apreciar en quina mesura les dades estadístiques presentades són vàlides i es poden prendre en consideració³⁹⁹, és a dir, apreciar que no constitueixen l'expressió de fenòmens merament fortuïts o conjunturals. També estableix la possibilitat d'apreciar la concurrència de discriminació si les dades estadístiques mostren una diferència menys important però persistent i relativament constant durant un llarg període de temps⁴⁰⁰. En l'assumpte Violeta Villar Láiz, va determinar que l'existència d'un desavantatge particular podria quedar demostrat provant que una normativa o disposició afecta negativament una proporció significativament més alta de persones d'un sexe que de l'altre⁴⁰¹.

Tot i així, no s'ha fixat per part del TJUE ni del TEDH un criteri únic sobre quin percentatge de persones negativament afectades per la disposició, criteri o pràctica aparentment neutra ha de pertànyer al grup desavantajat, ja que tendeixen a analitzar la situació de manera específica⁴⁰². Tampoc s'ha establert el tant per cent de persones amb el mateix atribut que han de resultar negativament afectades en comparació amb les altres que no comparteixen aquest atribut i que es troben en el grup desavantajat (exemple de

³⁹⁹TJCE, sentència de 9 de febrer de 1999, C-167/97, Regina contra Secretary of State for Employment, ex parte Nicole Seymour-Smith y Laura Pérez, par. 61 i 62.

⁴⁰⁰*Ídem*, par. 61; TJUE (Sala Tercera), sentència de 3 d'octubre de 2019, C- 274/18, Mínoo Schuch-Ghannadan contra Medizinische Universität Wien. En aquest assumpte estableix el TJUE que es pot demostrar la discriminació indirecta per qualsevol mitjà, inclús a partir de dades estadístiques (par. 46), donant la doctrina un pas més enllà i conclouent que es pot fonamentar la discriminació indirecta en base a dades estadístiques generals relatives al mercat de treball en l'estat membre del qual es tracti, en el supòsit que no s'espera que la persona interessada presenti dades més concretes relatives al grup de treballadors pertinent, per ser difícilment accessible o inclús no estar disponible (par. 56). Com a exemple de l'ús de proves estadístiques en òrgans jurisdiccionals nacionals, vegeu, en el cas espanyol, per exemple: Tribunal Constitucional (Sala Segunda), sentència 128/1987, de 16 de julio de 1987, FJ 5; Tribunal Constitucional (Pleno), sentència 253/2004, de 22 de diciembre de 2004, FJ 8. En els dos casos s'empra l'enquesta de Població activa per verificar, respectivament, la major taxa d'ocupació masculina i la feminització del treball a temps parcial. Més recentment vegeu: Tribunal Constitucional (Pleno), sentència 91/2019, de 3 de julio de 2019, FJ 10. En aquest cas es tracta de "*la consulta de estadísticas actuales en el mercado laboral*".

⁴⁰¹TJUE (Sala Tercera), sentència de 8 de maig de 2019, C-161/18, Violeta Villar Láiz contra Instituto Nacional de la Seguridad Social (INSS), par. 38.

⁴⁰²Vegeu, per exemple: TEDH (Gran Sala), sentència de 13 de novembre de 2007, 57325/00, D.H. i altres contra la República Txeca. En aquest assumpte, el TEDH va apreciar per primera vegada en la seva jurisprudència l'existència de discriminació indirecta al entendre que les proves psicològiques que s'empraven per determinar la intel·ligència i la idoneïtat -a la que se sotmetia a tots els alumnes per igual-, per avaluar si havien de ser apartats de l'educació general i adscrits a escoles dissenyades per nens amb necessitats especials causaven discriminació indirecta per raó d'ètnia, ja que entre el 80% i el 90% dels infants d'ètnia gitana acabaven fora del sistema general com a conseqüència d'aquestes proves. El tribunal va prendre aquesta decisió perquè va considerar desproporcionada la presència de menors d'ètnia gitana als centres escolars per infants amb necessitats especials ja que representaven un 50% dels infants dels centres quan només un 2% dels menors txecs eren d'ètnia gitana.

persones treballadores a temps parcial) per considerar que es dona un cas de discriminació indirecta⁴⁰³. És més, no ha establert requisits consistents per avaluar la fiabilitat, rellevància i importància de la prova estadística⁴⁰⁴. Sovint es delega l'establiment d'aquests requisits als tribunals nacionals, establint-se uns estàndards de prova divergents a la UE⁴⁰⁵. Les qüestions típicament es responen a escala nacional, regional o cas per cas segons factors locals, regionals o ambientals. Les lleis nacionals rarament defineixen tests, estàndards o mètriques de disparitat il·legal, o si ho fan, divergeixen entre els estats membres, fent que consegüentment, l'equitat contextual hagi dut a una protecció fragmentada⁴⁰⁶. També, com apunta l'Advocat General Léger en relació amb l'assumpte Inge Nolte, requerir estadístiques podria donar com a resultat una batalla de nombres que implícitament afavoreix a les persones demandants o demandades, segons qui és capaç de produir estadístiques més convincents⁴⁰⁷. És rellevant llavors com, estadístiques en relació amb assumptes complexos, per exemple en relació amb la discapacitat, l'orientació sexual o l'origen ètnic, moltes vegades directament ni estan disponibles⁴⁰⁸. En aquest sentit, en general, els tribunals nacionals i europeus i la doctrina són reticents a confiar en evidències estadístiques per provar les reclamacions per discriminació. Les estadístiques sovint no estan disponibles ni es veuen com confiabls. Sovint es prefereixen evidències intuïtives com el coneixement comú, fets obvis o conviccions⁴⁰⁹.

Finalment, existeix la problemàtica relativa a l'obtenció de dades que puguin concretar-se en una prova estadística per part de les persones que s'han vist discriminades. L'accés a informació sobre la decisió automatitzada al qual es refereix l'RGPD, com veurem en l'apartat 3 del capítol 3, és individual, de la mateixa manera que el dret a l'accés a les dades pròpies, ja que es protegeixen les dades personals d'altres persones. Aquella persona que s'ha vist discriminada no podrà accedir a les dades d'altres perquè estaran protegides per la legislació de protecció de dades⁴¹⁰. Inclús malgrat ser lògic pensar que

⁴⁰³SORIANO ARNANZ, A., "Decisiones automatizadas y discriminación...", *op. cit.*, 2021, pàg. 20.

⁴⁰⁴ WACHTER, S., MITTELSTADT, B. i RUSSELL, C., "Why fairness cannot be automated...", *op. cit.*, 2021, pàg. 18-20.

⁴⁰⁵TOBLER, C., *Limits and potential of the concept of indirect discrimination*, *op. cit.*, 2008, pàg. 41.

⁴⁰⁶WACHTER, S., MITTELSTADT, B. i RUSSELL, C., "Why fairness cannot be automated...", *op. cit.*, 2021, pàg. 45.

⁴⁰⁷Conclusiones del abogado general M. P. Léger, presentadas el 31 de mayo de 1995, en el asunto C-317/93, Inge Nolte contra Landesversicherungsanstalt Hannover, par. 53.

⁴⁰⁸WACHTER, S., MITTELSTADT, B. i RUSSELL, C., "Why fairness cannot be automated...", *op. cit.*, 2021, pàg. 34.

⁴⁰⁹ *Idem*, pàg. 65.

⁴¹⁰ RIVAS VALLEJO, P., "Discriminación algorítmica: detección, prevención y tutela", *op. cit.*, 2021, pàg. 48.

es tindran moltes proves estadístiques donat que s'analitzen dades massivament⁴¹¹, si la persona perjudicada no sap explotar-les no podran ser útils a l'hora provar la discriminació. Als Estats Units d'Amèrica del Nord, existeix la norma dels quatre cinquens que es basa en què una taxa de selecció per qualsevol sexe, raça o grup ètnic que és inferior a quatre cinquens de la taxa de selecció del grup amb la major taxa s'ha de considerar generalment com a evidència d'impacte advers⁴¹²(és a dir, discriminació indirecta). Aquesta regla adaptada a la discriminació algorítmica, suposaria reunir les estadístiques requerides per demostrar que una política és indirectament discriminatòria, procés que és sovint costos i difícil⁴¹³—com ja hem apuntat-, imposant una càrrega tècnica i financera a la persona demandant⁴¹⁴.

3.2.3. Justificació objectiva

La discriminació indirecta és un concepte basat en els efectes (mesura neutral que té efectes nocius) i en la regla de la raó⁴¹⁵ (noció adaptable a la persona demandada per evidenciar que la mesura és objectivament justificable).

Les justificacions de discriminació indirecta no són *numerus clausus*, contràriament als casos de discriminació directa⁴¹⁶. D'una banda, des d'una posició crítica amb el fet que existeixi la possibilitat de justificar objectivament els efectes discriminatoris d'un criteri, pràctica o disposició, Añón Roig estableix que el fet que hi hagi grups de població que hagin estat menysvalorats en la història i en conseqüència no hagin format part en els processos de formació de normes (p. e. dones), interpel·la directament a la visió dominant de la discriminació indirecta que sosté que aquesta pot justificar-se “objectivament”⁴¹⁷. El fet que la discriminació en els tractes neutres resulti difusa, i per tant, més complicada

⁴¹¹XENIDIS, R. i SENDEN, L., “EU Non-discrimination law...”, *op. cit.*, 2020, pàg. 171.

⁴¹²Uniform Guidelines on Employment Selection Procedures adopted by the United States of North America Equal Employment Opportunity Commission, Civil Service Commission, Department of Labor, and Department of Justice, 1978, par. 1607.4 (D).

⁴¹³TOBLER, C., *Indirect discrimination under Directives 2000/43 and 2000/78*, *op. cit.*, 2022, pàg. 99.

⁴¹⁴PÁEZ, A., “Negligent Algorithmic Discrimination”, *op. cit.*, 2021, pàg. 24 i 25.

⁴¹⁵En anglès “rule of reason”.

⁴¹⁶Cal puntualitzar aquí que a diferència de la discriminació directa on el TJUE estableix justificacions taxades i, en canvi, el TEDH té un sistema de justificació general objectiva; en el cas de la discriminació indirecta, tant el TJUE com el TEDH estableixen un sistema de justificació general objectiva.

⁴¹⁷ AÑÓN ROIG, M.J., “Desigualdades algorítmicas...”, *op. cit.*, 2022, pàg. 42.

d'identificar, no significa que hagi de poder-se justificar “objectivament”⁴¹⁸. D'altra banda, com planteja Maliszewska-Nienartowicz, es pot fer la lectura que el fet que la discriminació indirecta normalment no estarà relacionada amb la intenció de tracte desigual, que pot ocasionar un efecte discriminatori fins i tot accidentalment, suposa una raó suficient per a una gamma més àmplia de possibles motius de justificació⁴¹⁹. En aquest treball, i de manera congruent amb el que hem vingut defensant fins aquest punt, ens inclinem per la primera tesi segons la qual la discriminació indirecta no hauria de poder-se justificar objectivament.

Tanmateix, la normativa sí que permet la justificació objectiva. Dins d'aquests paràmetres, la justificació de la discriminació indirecta es basa en provar que la mesura discriminatòria persegueix una finalitat legítima, permetent al tribunal realitzar un examen de proporcionalitat a través del qual es ponderin els interessos en joc (adequació, necessitat i proporcionalitat en sentit estricte). No es passarà l'examen de proporcionalitat si la finalitat pot ser assolida a través d'una mesura diferent que sigui menys discriminatòria o no discriminatòria. En l'àmbit de la IA, tal com s'estableix en el document "*Directrices éticas para una IA fiable*", creades pel Grup d'Experts d'Alt Nivell sobre la Intel·ligència Artificial⁴²⁰, la IA hauria de ser usada per millorar el benestar col·lectiu i individual. Tot i així aquest concepte és bastant elàstic, i molts algorismes poden ser acceptats com a legítims sota aquesta base⁴²¹.

Es considera particularment conflictiu quan la situació contraposa dos drets humans, per exemple, si la presumpta persona discriminadora justifica en base a la llibertat religiosa una acció que involucra discriminació per raó de sexe. En general, el TJUE ha donat per vàlides pràcticament totes aquelles finalitats que no suposin un resultat discriminatori o que amaguin l'objectiu de discriminar de manera evident⁴²². Ens referim a “pràcticament totes les finalitats” ja que hi ha certs casos particulars, com per exemple en relació a les finalitats legítimes en el sector públic, en què el TJUE ha determinat que les

⁴¹⁸ BARRÈRE, M.A., “La igualdad de género desde el activismo de las profesiones jurídicas”, *Revista Vasca de Administración Pública*, vol. 1, núm. especial 99-100, 2014, pàg. 522 i 523. L'autora puntualitza que si per objectivitat es vol entendre neutralitat, ni la legitimitat (referida a la finalitat), ni l'adequació i necessitat (referida als mitjans), responen a paràmetres objectius.

⁴¹⁹ MALISZEWSKA-NIENARTOWICZ, J., “Direct and Indirect Discrimination...”, *op. cit.*, 2014, pàg. 54.

⁴²⁰ GRUPO DE EXPERTOS DE ALTO NIVEL SOBRE INTELIGENCIA ARTIFICIAL, *Directrices éticas para una IA fiable*, *op. cit.*, 2019, considerando 47.

⁴²¹ ALLEN, R. i MASTERS, D., “Artificial Intelligence: the right...”, *op. cit.*, pàg. 595.

⁴²² TOBLER, C., *Limits and potential of the concept of indirect discrimination*, *op. cit.*, 2008, pàg. 43; TOBLER, C., *Indirect discrimination under Directives 2000/43 and 2000/78*, *op. cit.*, 2022, pàg. 67.

consideracions de caràcter pressupostari no poden en si mateixes suposar una finalitat legítima que justifiqui una mesura indirectament discriminatòria⁴²³.

Presentem seguidament dos exemples en els quals el TJUE ha considerat els interessos econòmics com a objectius legítims vàlids dins el sector privat tot i comportar un clar tracte discriminatori envers un grup determinat de persones, per evidenciar com d'oberta és la possibilitat de justificació de la discriminació indirecta.

En el cas Achbita, el tribunal va establir que, en aplicació de l'article 16 de la CDFUE, s'havia de considerar legítima la voluntat de l'empresari de seguir un règim de neutralitat política, filosòfica o religiosa en les relacions amb els clients tant en el sector públic com en el sector privat malgrat el règim creava un desavantatge a les dones musulmanes ja que no podien portar hiyab⁴²⁴. En el mateix sentit, en l'assumpte Bilka-Kaufhaus, el tribunal va establir que conforme l'article 119 del Tractat de Roma⁴²⁵, es podia justificar l'adopció d'una política salarial que impliqués l'exclusió de les persones treballadores a temps parcial del règim de pensions d'empresa, amb independència del seu sexe, al·legant que es pretenia donar feina al menor nombre possible de persones treballadores a temps parcial⁴²⁶. En tot cas, sempre s'havia de constatar que els mitjans escollits per aconseguir aquest objectiu responien a una exigència efectiva de l'empresa, i eren adequats per aconseguir l'objectiu del qual es tractava i eren necessaris per a tal finalitat⁴²⁷.

En definitiva, una de les principals preocupacions relatives a que la discriminació algorítmica recaigui en la discriminació indirecta és que les possibilitats de justificar la conducta no estan taxades, i és més, el TJUE ha tendit, com hem vist, a acceptar-ne totes aquelles que, o bé no suposin un resultat discriminatori, o bé no amaguin l'objectiu de discriminar de manera evident. En el cas de la discriminació algorítmica, la tendència de les persones a confiar en la neutralitat dels sistemes automatitzats augmenta la possibilitat d'acceptació dels biaixos com a necessaris, i pot fer que s'arribi a concloure que superen

⁴²³Vegeu, per exemple: TJUE (Gran Sala), sentència d'11 de novembre de 2014, C-530/13, Leopold Schmitzer contra Bundesministerin für Inneres, par. 41.

⁴²⁴TJUE (Gran Sala), sentència de 14 de març de 2017, C-157/15, Samira Achbita i Centrum voor gelijkheid van kansen en voor racismebestrijding contra G4S Secure Solutions NV, par. 38.

⁴²⁵ Aquest article establí que cada estat membre garantiria l'aplicació del principi d'igualtat de retribució entre les persones treballadores de gènere masculí i femení per a una mateixa feina, prohibint la discriminació salarial per raó de gènere.

⁴²⁶TJUE, sentència de 13 de maig de 1986, C-170/84, Bilka - Kaufhaus GmbH contra Karin Weber von Hartz, par. 37.

⁴²⁷*Ídem, ibidem.*

el test de proporcionalitat, ocasionant, per consegüent, que els tribunals acabin acceptant decisions automatitzades que perpetuïn la situació de desavantatge de grups oprimits⁴²⁸.

3.2.4. Mitjans adequats i necessaris

La jurisprudència del TJUE evidencia que trobar una finalitat que sigui considerada legítima per una justificació objectiva és menys difícil que provar que els mitjans escollits són apropiats i necessaris⁴²⁹. Quan ens referim als sistemes algorítmics, la seva adequació com a mesura per aconseguir una finalitat establerta segurament dependrà de l'existència d'un altre algoritme que pugui aconseguir el mateix sense tenir uns efectes tan discriminatoris, ja que no existeixen altres maneres actualment de processar la quantitat de dades que tracten els algoritmes en la mateixa quantitat de temps. Per tant, tenint en compte l'opacitat i complexitat de la presa de decisions mitjançant sistemes algorítmics, és poc probable que es pugui analitzar si totes les variables que es tenen en compte són necessàries per la presa de la decisió o si prescindint d'elles també es podria arribar al mateix resultat; si la base de dades emprada és suficientment representativa o conté biaixos; o si el pes que se li dona a cada variable és adequat, resultant més probable que el test d'adequació es faci en relació al sistema en el seu conjunt, i no separatament⁴³⁰.

La part demandada, en teoria, haurà d'acreditar que va valorar l'ús d'altres sistemes i bases de dades i que no existia un algoritme menys discriminatori, ni una base de dades més completa i menys esbiaixada, tot i així, s'haurà de tenir en compte també la capacitat tècnica i econòmica de la part demandada⁴³¹, fet que estableix un ampli marge d'apreciació. La persona demandant podrà succeir si mostra que la persona demandada podria haver utilitzat un sistema alternatiu amb resultats menys discriminatoris⁴³². En el

⁴²⁸ SORIANO ARNANZ, A., "Discriminación algorítmica: garantías...", *op. cit.*, 2022, pàg. 162.

⁴²⁹ TOBLER, C., *Limits and potential of the concept of indirect discrimination*, *op. cit.*, 2008, pàg. 43.

⁴³⁰ HACKER, P., "Teaching fairness to artificial intelligence...", *op. cit.*, 2018, pàg. 1161 i 1162.

⁴³¹ SORIANO ARNANZ, A., "Decisiones automatizadas y discriminación...", *op. cit.*, 2021, pàg. 25.

⁴³² SORIANO ARNANZ, A., "Decisiones automatizadas: problemas y soluciones jurídicas. Más allá de la protección de datos", *Revista de Derecho Público: Teoría y Método*, vol. 3, 2021, pàg. 120-121. L'autora planteja en aquest article que l'establiment d'un sistema de millors tècniques disponibles facilitaria l'enjudiciament de decisions preses mitjançant sistemes algorítmics, ja que en el cas que s'estigués plantejant l'existència de discriminació algorítmica, el tribunal tindria la possibilitat de determinar si existeix un altre sistema algorítmic que, sense ser més costos ni menys precís, produeix resultats menys discriminatoris en el mateix context.

cas de les decisions automatitzades, això sembla difícil donada la natura opaca dels algoritmes que s'empren⁴³³. En els casos de contractació, per exemple, la persona demandant podrà evidenciar la possibilitat d'una alternativa que hagués pogut ésser menys discriminatòria i de la mateixa manera efectiva que la persona empresària hagués refusat emprar. En el cas que l'empresa utilitzi el seu propi model ja serà difícil provar que una alternativa millor existeix, però esdevé un problema irresoluble quan els models emprats són de terceres parts, i la persona de qui depèn finalment la decisió no té cap possibilitat de retocar el model. En aquest cas, a la persona que ha pres la decisió únicament podrà reclamar-se-li responsabilitat per haver contractat una empresa que empra un model discriminatori. Aquest seria l'exemple de persones empresàries que utilitzen Facebook o HireVue per contractar, ja que en no utilitzar un model propi, només se les podrà fer responsables d'haver contractat un sistema de selecció –que s'haurà de provar, en el cas que no estigui provat- que és discriminatori, i en aquest cas, serà molt fàcil per a la persona que ha pres la decisió eludir la seva responsabilitat⁴³⁴.

Com ja hem dit quant a l'equitat dels algoritmes i l'aparença que aquests són justos i precisos, es considera que sempre són el producte d'una societat basada en certes estructures de discriminació, amb biaixos imperceptibles que es manifesten en el tractament desigual dels grups⁴³⁵. S'ha de tenir en compte que ni la supra ni la infra representació de membres de la societat vulnerables ni el mecanisme a través del qual succeeix no són sempre evidents. Així que mentre teòricament una manera menys discriminatòria hauria d'existir sempre, provar-ho pot esdevenir difícil⁴³⁶.

En definitiva, el concepte de discriminació indirecta resulta en estàndards oberts que són sovint difícils d'aplicar⁴³⁷, llavors, malgrat doctrinalment es consideri que pot captar millor la discriminació algorítmica, també té certes deficiències. En aquest sentit també, considera Sáez Lara que l'aplicació mecànica de la doctrina de la discriminació indirecta falla perquè els models algorítmics no presenten els caràcters que s'identifiquen per l'aplicació correcta d'aquesta doctrina, atès que en comptes de proporcionar criteris

⁴³³PÁEZ, A., “Negligent Algorithmic Discrimination”, *op. cit.*, 2021, pàg. 20.

⁴³⁴*Ídem*, pàg. 27.

⁴³⁵ADAMS-PRASSL, J., BINNS, R. i KELLY-LYTH, A., “Directly Discriminatory Algorithms”, *op. cit.*, 2023, pàg. 152; FRIEDLER, A., SCHEIDEGGER, C. i VENKATASUBRAMANIAN, S., “The (Im)possibility of Fairness: Different Value Systems Require Different Mechanisms For Fair Decision Making”, *Communications of the ACM*, vol. 64, núm. 4, 2017, pàg. 139.

⁴³⁶BAROCAS, S. i SELBST, A. D., “Big Data’s Disparate...”, *op. cit.*, 2016, pàg. 710.

⁴³⁷ZUIDERVEEN, F., *Discrimination, artificial intelligence...*, *op. cit.*, 2018, pàg. 34.

d'avaluació que estan justificats per la seva rellevància, els models de dades involucren processos de decisió opacs, que descansen en correlacions inexplicables i els manquen justificacions clarament articulades⁴³⁸. S'afegeix la dificultat de provar un desavantatge particular en base a proves estadístiques (úniques que semblen factibles per provar discriminació algorítmica) i que el TJUE tendeix a acceptar aquelles justificacions que no suposin un resultat discriminatori de manera evident o amaguin l'objectiu de discriminar.

Per tant, el problema principal de la vessant indirecta de la discriminació algorítmica no serà provar que la tipologia de discriminació ha estat indirecta sinó aconseguir elements probatoris que constatin l'existència de discriminació. Així i tot, serà un pas menys que evidenciar discriminació directa, en la qual, segurament, una vegada constatada la discriminació, caldrà provar que aquesta ha estat directa.

Seguidament abordem perquè la presa de decisions mitjançant sistemes algorítmics és més procliu a ocasionar discriminació indirecta que discriminació directa.

3.3. Per què la discriminació algorítmica es tradueix sobretot en discriminacions indirectes?

El TJUE, quan analitza casos de discriminació, normalment delega al tribunal nacional la decisió de si es tracta d'un cas de discriminació directa o de discriminació indirecta⁴³⁹. Tanmateix, la delimitació entre els dos conceptes, planteja un desafiament particular pels tribunals nacionals⁴⁴⁰. Seguint a Tobler, la jurisprudència del TJUE actualment distingeix l'existència de discriminació directa o indirecta de la següent manera⁴⁴¹: es tractarà de discriminació directa quan, o bé s'ha fet un ús exprés de la categoria protegida per la

⁴³⁸SÁEZ LARA, C., "El algoritmo como protagonista de la relación laboral...", *op. cit.*, 2020, pàg. 48 i 49.

⁴³⁹ Vegeu, per exemple: TJUE (Sala Segona), sentència de 21 de juliol de 2011, C-104/10, Patrick Kelly contra National University of Ireland (University College, Dublin), par. 31; TJUE (Sala Segona), sentència de 19 d'abril de 2012, C-415/10, Galina Meister contra Speech Design Carrier Systems GmbH, par. 37; TJUE (Gran Sala), sentència de 16 de juliol de 2015, C-83/14, CHEZ Razpredelenie Bulgaria AD contra Komisia za zashtita ot diskriminatsia, par. 79; TJUE (Gran Sala), sentència de 26 de gener de 2021, C-16/19, VL contra Szpital Kliniczny im. dra J. Babińskiego Samodzielny Publiczny Zakład Opieki Zdrowotnej w Krakowie, par. 60.

⁴⁴⁰ TOBLER, C., *Indirect discrimination under Directives 2000/43 and 2000/78*, *op. cit.*, 2022, pàg. 113.

⁴⁴¹ *Ídem*, pàg. 88.

legislació (vegeu, per exemple, el cas Mangold⁴⁴²), o bé es pot provar que un criteri s'ha escollit per raons relacionades amb una categoria protegida per la legislació, malgrat ser formalment neutre (vegeu, l'argumentació que fa el tribunal en el cas CHEZ⁴⁴³), o bé el criteri usat és inseparable o està inextricablement vinculat a una categoria protegida per la legislació (vegeu, per exemple, el cas Szpital Kliniczny⁴⁴⁴). Sembla que quan es tracta d'un criteri inseparable de la categoria protegida, el TJUE es focalitza en l'efecte exclouent del criteri emprat, o bé totes les persones que es veuen perjudicades pel criteri formen part del mateix grup protegit (vegeu per exemple, el cas Nikoloudi⁴⁴⁵), o bé el grup protegit inclou persones que estan completament excloses, en el sentit que mai podran complir el criteri en qüestió (vegeu, per exemple, el cas Hay⁴⁴⁶). Per contra, es tractarà de discriminació indirecta quan s'emprï qualsevol altre criteri formalment neutre, on tant el grup avantatjat com el grup desavantatjat siguin heterogenis (vegeu, per exemple, el cas Bilka-Kaufhaus⁴⁴⁷).

Però, com s'adapta la discriminació mitjançant sistemes algorítmics a aquests dos conceptes? Rivas Vallejo posa l'exemple dels sistemes algorítmics de contractació. Estaríem davant de discriminació directa si l'empresa (la que crea el sistema per utilitzar-lo directament o la que el comercialitza per a altres empreses) trasllada o introdueix paràmetres de biaix en l'algoritme (aquest es construeix en base a ordres que persegueixen un resultat concret i aquest es defineix per qui ordena la seva programació).

En canvi, estaríem davant de discriminació indirecta en el cas d'un algoritme d'aprenentatge automàtic que funciona de manera opaca, simplement valora tots els elements que han posat a la seva disposició per arribar a la "millor decisió", estant aquesta última esbiaixada però de manera aliena a una motivació de qui ha programat l'algoritme.

⁴⁴² TJCE (Gran Sala), sentència de 22 de novembre de 2005, C-144/04, Weiner Mangold contra Rüdiger Helm, par. 57.

⁴⁴³ TJUE (Gran Sala), sentència de 16 de juliol de 2015, C-83/14, CHEZ Razpredelenie Bulgaria AD contra Komisia za zashtita ot diskriminatsia, par. 91.

⁴⁴⁴ TJUE (Gran Sala), sentència de 26 de gener de 2021, C-16/19, VL contra Szpital Kliniczny im. dra J. Babińskiego Samodzielny Publiczny Zakład Opieki Zdrowotnej w Krakowie, par. 48.

⁴⁴⁵ TJCE, sentència de 10 de març de 2005, C-196/02, Vasiliki Nikoloudi contra Organismos Tilepikoinonion Ellados AE, par. 40. En el mateix sentit, vegeu també: TJUE, sentència de 18 de novembre de 2010, C-356/09, Pensionsversicherungsanstalt contra Christine Kleist, par. 42 i; TJUE (Gran Sala), sentència de 12 d'octubre de 2010, C-499/08, Ingeniørforeningen i Danmark, acting on behalf of Ole Andersen contra Region Syddanmark, par. 23 i 24.

⁴⁴⁶ TJUE, sentència de 12 de desembre de 2013, C-267/12, Frédéric Hay contra Crédit agricole mutuel de Charente-Maritime et des Deux-Sèvres, par. 44.

⁴⁴⁷ TJUE, sentència de 22 de novembre de 2012, C-385/11, Isabel Elbal Moreno e Instituto Nacional de la Seguridad Social (INSS), Tesorería General de la Seguridad Social (TGSS), par. 38.

Si, igualment, l'empresa pren de referència un patró històric discriminatori per arribar a la “millor decisió”, és a dir, els antecedents de biaix de la pròpia empresa, estaria duent a terme un tracte obertament desfavorable cap a aquelles persones dotades d'un atribut protegit (per exemple, sexe o origen ètnic) que s'han vist perjudicades per la política de contractació en el passat, de manera que la discriminació també seria directa⁴⁴⁸.

Aquí se'ns presenten diversos casos:

1. Si s'han donat al sistema algorítmic directrius concretes i evidents per tractar desfavorablement a un grup determinat de població que comparteix un atribut protegit, estarem davant de discriminació directa.
2. Si el sistema algorítmic és d'aprenentatge automàtic i en base als patrons, tendències o regles socials que ha determinat en les dades, simplement valora tots els elements que s'han posat a la seva disposició per prendre la “millor decisió”, estarem davant d'un cas de discriminació indirecta, atès que l'algoritme haurà tingut en compte altres factors que no són únicament l'atribut protegit (i que no són equivalents a aquest) per prendre la decisió, però el resultat haurà estat més desfavorable per un grup de persones que comparteixen un atribut protegit determinat⁴⁴⁹. No obstant això, aquests altres factors no poden ser “equivalents a l'atribut protegit”, ja que si són equivalents podríem estar davant d'un cas de discriminació directa.
3. Si el sistema algorítmic és d'aprenentatge automàtic i per arribar a la “millor decisió” emprades dades que venen donades en base a un patró històric discriminatori (que es pot concretar en les dades històriques de contractació d'una empresa, per exemple), llavors es tracta de discriminació directa, atès que d'entrada qui empra

⁴⁴⁸ RIVAS VALLEJO, P., “Sesgos de género en el uso...”, *op. cit.*, 2022, pàg. 67-68. L'autora fa aquest plantejament per concloure que en l'àmbit laboral, la persona empresària continua essent responsable de l'acció discriminatòria, essent determinant l'actitud de l'empresa enfront el possible efecte pervers de l'algoritme, és a dir, el seu consentiment per validar el biaix discriminatori que reproduceix, confirmant la proposta del model matemàtic.

⁴⁴⁹ Ens podríem preguntar aquí si en tots els casos que s'emprin sistemes algorítmics que incloguin dades de comportament hi haurà discriminació atès que els resultats mostraran l'estat actual del món, fruit de relacions de dominació i subordinació. Així que, tothom que empri sistemes algorítmics de presa de decisió emprarà bases de dades discriminatòries (que reflecteixen l'estat passat o actual del món) que donaran resultats indefectiblement indissociables del tracte discriminatori que reben les persones que ostenten atributs protegits.

l'algoritme ja sabrà que els resultats que reveli seran discriminatoris i que el tracte ja era desigual des del començament.

Sembla ser, llavors, que en l'estat actual de la qüestió, la característica principal que distingeix la discriminació directa de la discriminació indirecta és el coneixement per part de la persona que pren la decisió sobre la relació entre l'atribut protegit i el resultat.

Cal tenir en compte en la distinció de discriminació directa i indirecta plantejada per la jurisprudència del TJUE, que tota la construcció dels conceptes està creada en base a la discriminació per un atribut protegit, per tant, no s'aborda una possible multiplicitat d'atributs protegits que intervinguin en la decisió.

En base a les distincions fetes per Rivas Vallejo, podem identificar dos elements que dificulten la distinció entre discriminació directa i indirecta (és a dir, la distinció entre si hi havia, o no, coneixement sobre la relació entre l'atribut protegit i el resultat). En primer lloc, la possibilitat de fer ús d'indicadors molt propers a l'atribut protegit que puguin ser, al cap i a la fi, equivalents a l'atribut protegit, però que no es pugui determinar que són equivalents (3.3.1); i en segon lloc, el caràcter subreptici de la discriminació algorítmica, que pot causar que malgrat les discriminacions siguin directes, sigui impossible provar-ho (3.3.2).

3.3.1. Factors equivalents a l'atribut protegit

El funcionament dels sistemes algorítmics i la seva opacitat generen, entre la distinció tradicionalment establerta de discriminació directa i discriminació indirecta, un repte de classificació, donat que serà cada cop més difícil identificar tractament diferencial basat en atributs protegits en el context de les operacions algorítmiques⁴⁵⁰. En aquest sentit, conclouen Barocas i Selbst que la discriminació directa i la discriminació indirecta esdevenen essencialment el mateix des d'una perspectiva probatòria⁴⁵¹.

⁴⁵⁰GERARDS, J. i XENIDIS, R., *Algorithmic Discrimination in Europe...*, op. cit., 2021, pàg. 76.

⁴⁵¹BAROCAS, S. i SELBST, A. D., "Big Data's Disparate...", op. cit., 2016, pàg. 713.

La qüestió resideix en si s'aborden els motius protegits des d'una perspectiva restrictiva o àmplia. Més enllà de la identificació de motius protegits en les decisions preses mitjançant algorismes, existeixen les correlacions i els indicadors que tenen rellevància en relació a la definició mateixa d'atribut protegit, i qüestionen com d'estretament o àmpliament s'haurien d'entendre. En el cas Szpital Kliniczny, el tribunal estableix que existeix discriminació directa per raó de discapacitat, quan el tractament desfavorable es basa en un criteri inextricablement vinculat a la discapacitat⁴⁵². Usualment hi ha una clara superposició entre el motiu protegit i els seus indicadors, és a dir, que l'ús de l'indicador cobreix quasi exactament el mateix grup de persones que cobriria el motiu protegit⁴⁵³.

Si se'ls dona una interpretació extensiva, els indicadors més rellevants també haurien d'estar coberts per la protecció de discriminació directa, ja que malgrat no ser exactament l'atribut es produeix una superposició del significat quasi completa. L'exemple clàssic és la discriminació basada en l'embaràs, que està clarament relacionada amb la discriminació per raó de sexe ja que embaràs es un indicador de dona. El TJUE ha tractat la discriminació per embaràs com a discriminació directa⁴⁵⁴. Un altre bon exemple és tenir un passaport estranger que és un indicador clar de tenir una nacionalitat diferent.

Contràriament, si s'entén l'atribut protegit de manera restrictiva, només quan es discrimini en base a aquest estarem parlant de discriminació directa, i tota la discriminació que provingui dels indicadors o correlacions sobre aquest atribut es considerarà discriminació indirecta. Per tant, un repte important que sorgeix amb la discriminació algorítmica és descobrir quins indicadors arriben a tenir un grau tan gran de superposició amb el motiu protegit que poden ser vistos com la mateixa cosa⁴⁵⁵. S'ha de qüestionar llavors si és necessari un 100% de superposició o estadísticament també s'admet un indicador tingui un 90% o 80% de superposició amb un atribut protegit; i si aquest tractament desigual pot seguir-se considerant discriminació directa o s'ha de considerar

⁴⁵² TJUE (Gran Sala), sentència de 26 de gener de 2021, C-16/19, VL contra Szpital Kliniczny im. dra J. Babińskiego Samodzielny Publiczny Zakład Opieki Zdrowotnej w Krakowie, par. 48.

⁴⁵³ GERARDS, J. i XENIDIS, R., *Algorithmic Discrimination in Europe...*, *op. cit.*, 2021, pàg. 76.

⁴⁵⁴ Vegeu: TJCE, sentència de 8 de novembre de 1990, C-177/88, Elisabeth Johanna Pacifica Dekker contra Stichting Vormingscentrum voor Jong Volwassenen (VJV-Centrum) Plus. El TJUE en aquest assumpte va determinar que rebutjar una única candidata a un lloc de treball a causa d'estar embarassada era discriminació directa per raó de sexe, ja que només es pot donar el cas si la persona candidata és dona (par. 12). Aquesta situació és difícilment comparable en el cas que la decisió sigui presa per un algorisme al contingut del qual no es pugui accedir.

⁴⁵⁵ ADAMS-PRASSL, J., BINNS, R. i KELLY-LYTH, A., "Directly Discriminatory Algorithms", *op. cit.*, 2023, pàg. 171.

indirecta o discriminació per associació amb el grup que es caracteritza amb l'atribut protegit⁴⁵⁶.

Cercar tractament diferencial basat en atributs sensibles pot esdevenir molt complicat i pot ocasionar que al final, en no localitzar si la discriminació prové directament d'un atribut sensible (sexe, origen ètnic, etc.) o d'un indicador (embaràs, passaport estranger, etc.), s'acabi sense poder determinar exactament si es tracta de discriminació directa o indirecta, fent que al final sigui la discriminació indirecta la que actuï com a refugi per captar la discriminació algorítmica. Això pot suposar, com ja hem dit, la reducció de la seguretat jurídica si significa que es podrà aplicar el sistema de justificacions objectives obertes en comptes del conjunt *numerus clausus* de justificacions disponibles per la discriminació directa⁴⁵⁷. Per detectar la discriminació directa en els sistemes algorítmics s'ha proposat que es pugui experimentar amb els sistemes⁴⁵⁸, i així tenir una eina per intentar evitar que la discriminació indirecta actuï com a refugi de la discriminació algorítmica.

3.3.2. Altres causes que motiven que la discriminació algorítmica recaigui en la discriminació indirecta

Com hem vist, si bé és cert que es pot donar la discriminació directa en els sistemes algorítmics⁴⁵⁹, majoritàriament, s'enquadrarà en el concepte de discriminació indirecta⁴⁶⁰, a no ser que es posi de manifest que el funcionament de l'algoritme (ja sigui a través de

⁴⁵⁶ GERARDS, J. i XENIDIS, R., *Algorithmic Discrimination in Europe...*, *op. cit.*, 2021, pàg. 64.

⁴⁵⁷ *Ídem*, pàg. 9.

⁴⁵⁸ SORIANO ARNANZ, A., "La aplicación del marco jurídico europeo en materia de igualdad y no discriminación al uso de aplicaciones de Inteligencia Artificial", a BONORINO RAMÍREZ, P. R., FERNÁNDEZ ACEVEDO, R. i VALCÁRCEL FERNÁNDEZ, P., (dirs.), *Nuevas normatividades: inteligencia artificial, derecho y genero*, Navarra, Aranzadi, 2021, pàg. 81.

⁴⁵⁹ ADAMS-PRASSL, J., BINNS, R. i KELLY-LYTH, A., "Directly Discriminatory Algorithms", *op. cit.*, 2023, pàg. 154-164.

⁴⁶⁰Vegeu: HACKER, P., "Teaching fairness to artificial intelligence...", *op. cit.*, 2018, pàg. 1143; ZUIDERVEEN, F., "Strengthening legal protection against discrimination..." *op. cit.*, 2020, pàg. 1578; BAROCAS, S. i SELBST, A. D., "Big Data's Disparate...", *op. cit.*, 2016, pàg. 671; ZARSKY, T., "An analytic challenge: discrimination theory in the age of predictive analytics", *A Journal of Law and Policy for the Information Society*, vol. 14, núm. 1, 2017, pàg. 23.

la programació, ja sigui a través de bases de dades esbiaixades), provoca una diferència de tracte per raons lligades a un atribut protegit.

El tractament de les dades i la seva categorització per part dels algoritmes pot no ser cognoscible per la persona i per tant, seria impossible determinar on hi ha la discriminació directa⁴⁶¹. A més, l'ús de variables i categorització de dades en els algoritmes d'aprenentatge automàtic està en constant evolució mentre l'algoritme aprèn. Com aquestes categories no són estàtiques, serà difícil de saber si estan relacionades amb els atributs protegits. S'haurà de mirar el model algorítmic i la manera com el model estadístic usat tracta les dades disponibles durant el temps per trobar com actua el tractament desfavorable, una situació que pot ser difícil d'acomplir donats els problemes d'accessibilitat⁴⁶².

Atesa la prohibició d'emprar atributs protegits en els sistemes algorítmics (llevat que s'apliqui l'excepció de l'article 10.5 de la PRIA, tal com abordarem en l'apartat 1.3 del capítol VI, que estableix que en la mesura en què sigui estrictament necessari per garantir la vigilància, detecció i correcció dels biaixos associats al sistema d'IA de risc alt, les persones proveïdores dels sistemes podran incloure en el tractament atributs protegits), la discriminació directa podria ser menys probable en la discriminació algorítmica en comparació amb la discriminació duta a terme per éssers humans⁴⁶³. No s'empraran directament, no només pel fet d'evitar que es pugui detectar la discriminació més fàcilment que a través de la discriminació indirecta, que és més subtil, sinó que les persones que desenvolupin els sistemes algorítmics, podrien no introduir en l'algoritme les dades relacionades directament amb atributs protegits⁴⁶⁴ per no ocasionar una disminució de la precisió algorítmica, que incrementa com més correlacions diverses es poden fer⁴⁶⁵.

A més, els sistemes algorítmics, en aplicar patrons repetitius, tendències o regles que expliquin el comportament de dades en un determinat context, perpetuen accions discriminatòries no intencionals⁴⁶⁶ -més fàcils de passar per alt-, que les persones hem anat adoptant a causa de viure en una societat desigual i marcada per les discriminacions

⁴⁶¹ LEESE, M., "The New Profiling: Algorithms, Black Boxes...", *op. cit.*, 2014, pàg. 494.

⁴⁶² GERARDS, J. i XENIDIS, R., *Algorithmic Discrimination in Europe...*, *op. cit.*, 2021, pàg. 69.

⁴⁶³ *Ídem, ibidem.*

⁴⁶⁴ SCHREURS, W., "Cogitas, Ergo Sum. The role of data protection...", *op. cit.*, 2008, pàg. 260.

⁴⁶⁵ HACKER, P., "Teaching fairness to artificial intelligence...", *op. cit.*, 2018, pàg. 1152 i 1161.

⁴⁶⁶ BAROCAS, S. i SELBST, A. D., "Big Data's Disparate...", *op. cit.*, 2016, pàg. 693.

sistèmiques i estructurals però de les quals molt probablement no en som conscients (típicament enquadrades en el concepte de discriminació indirecta). Moltes de les dades incloses en entrenar un algoritme i les correlacions que s'efectuen són el resultat dels prejudicis implícits de la persona qui els entrena, que al seu torn pot ser un reflex dels estereotips culturals que prevalen en el seu entorn. En els casos en què la persona que entrena l'algoritme emprava un model predictiu discriminatori, les legislacions antidiscriminació són fàcils d'eludir fent ús d'indicadors que representin ètnia, religió o sexe. Provar que la disfressa dels indicadors és intencionada és difícil, si no impossible, ja que qui ho ha fet sempre pot argumentar que no era conscient que això s'estava donant (vegeu apartat 2.2.1 del capítol I). D'aquesta manera, aquests mètodes de discriminació intencional semblaran, a tots els efectes, idèntics a la discriminació no intencional que pot resultar de l'anàlisi de dades⁴⁶⁷.

Per un tribunal, és difícil -si no impossible- fer una retrospectiva i reconstruir les nombroses avaluacions esbiaixades i percepcions que han resultat en la decisió adversa⁴⁶⁸, basada en biaixos implícits i inconscients que es donen quan alguna persona ha internalitzat un estereotip social de tal manera que aplica a una persona un tractament ostensiblement més desfavorable que a una altra.

Finalment, s'ha posat atenció en què quan les obligacions legals augmenten, la discriminació directa és procliu a disminuir en la fase de desenvolupament dels algoritmes⁴⁶⁹.

A tall de recapitulació, la discriminació indirecta s'evidencia quan existeixen efectes desfavorables d'un criteri, disposició o pràctica aparentment igual per tothom, en grups que comparteixen atributs protegits per la legislació. Sembla més adequada per enquadrar la discriminació algorítmica, no tant perquè usualment no es produeixen *per se* discriminacions directes, sinó que aquestes, atesa l'opacitat dels sistemes algorítmics semblen impossibles de provar. Que el funcionament de la discriminació algorítmica s'enquadri millor en la discriminació indirecta suposa dos problemes. En primer lloc, que les justificacions de discriminació directa estan taxades en el dret de la UE, contràriament a aquelles de la discriminació indirecta, que són obertes. En el cas que recaigui tot el pes

⁴⁶⁷ PÁEZ, A., "Negligent Algorithmic Discrimination", *op. cit.*, 2021, pàg. 26.

⁴⁶⁸ BAROCAS, S. i SELBST, A. D., "Big Data's Disparate...", *op. cit.*, 2016, pàg. 698.

⁴⁶⁹ HACKER, P., "Teaching fairness to artificial intelligence...", *op. cit.*, 2018, pàg. 1152.

de la protecció de les discriminacions algorítmiques en la discriminació indirecta, la possibilitat de justificar-les pot ser molt més àmplia⁴⁷⁰. En segon lloc, i com tractem amb més precisió seguidament, la prova de discriminació indirecta pot ser molt difícil d'aconseguir en casos de discriminació algorítmica.

4. Deficiències respecte a la prova en el dret antidiscriminatori europeu per abordar la discriminació mitjançant sistemes algorítmics

És difícil establir mètriques d'equitat suficientment adequades i a la vegada estàtiques per ser codificades. La llei i la jurisprudència enquadren l'equitat i la discriminació com a conceptes fonamentalment contextuals. Al mateix temps, l'ús creixent d'algoritmes està pertorbant els procediments i solucions tradicionals per prevenir i investigar casos de discriminació⁴⁷¹. Definicions clau, estàndards de prova i percentatges comparatius estan definits en base a les circumstàncies del cas que inclouen qui es un grup desavantatjat i comparador adequat; què constitueix un desavantatge particular (econòmic, oportunitats perdudes...); la rellevància del dany causat per la discriminació (ha d'afectar a un gran nombre de persones, no ha de ser fortuït, normes menys estrictes si la discriminació ha estat constant en un llarg període de temps...); si s'ha de provar un dany concret o una amenaça potencial és suficient; o quin tipus d'evidències són admissibles davant el tribunal⁴⁷².

Emprar procediments judicials per fer complir els principis d'igualtat i no-discriminació en els algoritmes està ple de dificultats. En molts països manca el coneixement de la judicatura i de les persones que treballen en l'advocacia en relació a la legislació antidiscriminació⁴⁷³. Segons un seguit d'entrevistes fetes a persones expertes en normativa antidiscriminària d'estats membres de la UE, pot ser oferta poca protecció contra la discriminació a Grècia, Itàlia o Espanya, ja que en aquests països, difícilment es porten davant dels tribunals nacionals casos de discriminació i, si es fa, pot haver-hi llargs

⁴⁷⁰ GERARDS, J. i XENIDIS, R., *Algorithmic Discrimination in Europe...*, op. cit., 2021, pàg. 73.

⁴⁷¹ WACHTER, S., MITTELSTADT, B. i RUSSELL, C., "Why fairness cannot be automated...", op. cit., 2021, pàg. 64.

⁴⁷² *Ídem*, pàg. 64 i 65.

⁴⁷³ GERARDS, J. i XENIDIS, R., *Algorithmic Discrimination in Europe...*, op. cit., 2021, pàg. 119.

retards que fan que al final no ofereixin compensació a la víctima⁴⁷⁴. És més, no només per a les persones expertes pot ser difícil identificar la font de la discriminació algorítmica, sinó que qualsevol evidència de discriminació en el procés pot haver desaparegut quan l'algoritme ha “mutat”, “evolucionat”. També, malgrat les persones expertes puguin localitzar la font de discriminació, pot ser impossible fer-ho per part de les víctimes, la judicatura, les persones legisladores, etc.⁴⁷⁵

Cal llavors analitzar les dificultats que presenta el repartiment de la càrrega de la prova en els casos de discriminació algorítmica (4.1), l'establiment de la responsabilitat de la discriminació algorítmica (4.2) i finalment, l'absència de *tertium comparationis* (4.3).

4.1. Repartiment de la càrrega de la prova

La normativa antidiscriminària europea es basa en la litigació individual. La persona que presumptivament s'ha vist discriminada ha d'aportar indicis de discriminació *prima facie*. Quan es consideren establerts els indicis, s'inverteix la càrrega de la prova i és la persona presumptivament discriminadora qui ha de demostrar que no està realitzant cap conducta discriminatòria⁴⁷⁶. La inversió de la càrrega de la prova pot ocórrer quan la persona demandant ha establert una discriminació *prima facie* o quan el nexce causal entre

⁴⁷⁴ *Ídem*, pàg. 115.

⁴⁷⁵ *Ídem*, pàg. 74.

⁴⁷⁶ Art. 8 de la Directiva 2000/43/CE; art. 10 de la Directiva 2000/78/CE; art. 19 de la Directiva 2006/54/CE i; art. 9 de la Directiva 2004/113/CE. En l'ordenament jurídic espanyol (que fa una correcta adaptació de la normativa europea), la disposició de la càrrega de la prova en l'àmbit civil i contenciós-administratiu es regula en l'article 32 de la Llei 62/2003, que disposa el següent: “*En aquellos procesos del orden jurisdiccional civil y del orden jurisdiccional contencioso-administrativo en que de las alegaciones de la parte actora se deduzca la existencia de indicios fundados de discriminación por razón del origen racial o étnico de las personas, corresponderá al demandado la aportación de una justificación objetiva y razonable, suficientemente probada, de las medidas adoptadas y de su proporcionalidad*”. En l'àmbit contenciós administratiu també s'estableix aquest repartiment de la càrrega de la prova en l'article 30 de la Llei 15/2022. En l'àmbit social, es regula en l'article 96 de la Ley 62/2003, de 30 de diciembre, de medidas fiscales, administrativas y del orden social, i versa de la següent manera: “*En aquellos procesos en que de las alegaciones de la parte actora se deduzca la existencia de indicios fundados de discriminación por razón de sexo, orientación o identidad sexual, origen racial o étnico, religión o convicciones, discapacidad, edad, acoso y en cualquier otro supuesto de vulneración de un derecho fundamental o libertad pública, corresponderá al demandado la aportación de una justificación objetiva y razonable, suficientemente probada, de las medidas adoptadas y de su proporcionalidad*”. També l'article 30 de la Llei 15/2022. El tercer apartat d'aquest article també precisa que el repartiment de la càrrega de la prova que s'estableix en aquests articles no s'aplica en els procediments penals, on preval el principi de presumpció d'innocència.

l'atribut protegit i el dany és probable⁴⁷⁷. Assumint que el canvi de la càrrega de la prova ocorre, en general hi ha dues maneres a través de les quals la persona demandada pot defensar-se: o bé refutant que existeixi un nexa causal entre els resultats discriminatoris i l'atribut protegit⁴⁷⁸; o bé reconeixent que els resultats discriminatoris han ocorregut però es pot donar una justificació, en base a un interès legítim, de manera necessària i proporcionada⁴⁷⁹.

La sentència Flauchère, dictada pel Tribunal de Cassació francès el 28 de març del 2000⁴⁸⁰, integra l'enfocament comunitari construint una teoria de l'accés a prova en matèria de discriminació sindical⁴⁸¹. En el cas, un treballador reclama que, com a conseqüència de la visió negativa per part de l'empresari de les seves activitats sindicals, hi ha hagut una dilació en el desenvolupament de la seva carrera. El Tribunal estableix que malgrat que correspon a la persona que presumptivament s'ha vist discriminada presentar els elements de fet susceptibles de caracteritzar una violació del principi d'igualtat, correspon a la persona que presumptivament ha causat la discriminació d'establir que la disparitat de tractament soferta per la persona està justificada per elements objectius aliens a qualsevol discriminació i que correspon al jutge verificar, en presència de la discriminació sindical invocada, les condicions en les quals la carrera de la persona presumptivament discriminada s'ha desenvolupat⁴⁸².

⁴⁷⁷ RINGELHEIM, J., "The Burden of Proof in Antidiscrimination Proceedings. A Focus on Belgium, France and Ireland", *European Equality Law Review*, núm. 2, 2019, pàg. 52.

⁴⁷⁸ També es pot emprar aquesta defensa en la discriminació directa.

⁴⁷⁹ TJUE, sentència de 13 de maig de 1986, C-170/84, Bilka - Kaufhaus GmbH contra Karin Weber von Hartz, par. 35.

⁴⁸⁰ Cour de cassation, décision 97-45258, du 28 mars 2000. Disponible a: https://www.courdecassation.fr/decision/6079b1a49ba5988459c52c35?search_api_fulltext=97-45258&op=Rechercher+sur+judilibre&date_du=2000-03-28&date_au=2000-03-28&judilibre_jurisdiction=cc&previousdecisionpage=&previousdecisionindex=&nextdecisionpage=&nextdecisionindex= Darrera visita el 2 de setembre de 2023.

⁴⁸¹ ESTEVE GARCÍA, F., "Las directivas europeas contra la discriminación racial y la creación de organismos especializados para promover la igualdad. Análisis comparativo de su transposición en España y en Francia", *Revista de Derecho Constitucional Europeo*, año 5, núm. 10, 2008, pàg. 18.

⁴⁸² La sentència de la Cour de cassation estableix que: "*Attendu, cependant, qu'il appartient au salarié syndicaliste qui se prétend lésé par une mesure discriminatoire de soumettre au juge les éléments de fait susceptibles de caractériser une atteinte au principe d'égalité de traitement et qu'il incombe à l'employeur, s'il conteste le caractère discriminatoire du traitement réservé au syndicaliste, d'établir que la disparité de situation constatée est justifiée par des éléments objectifs, étrangers à toute discrimination fondée sur l'appartenance à un syndicat ; Qu'en statuant comme elle l'a fait, alors que, d'une part, si le juge n'a pas à se substituer à l'employeur, il lui appartient de vérifier, en présence d'une discrimination syndicale invoquée, les conditions dans lesquelles la carrière des intéressés s'est déroulée, et alors, d'autre part, que la preuve de la discrimination n'incombait pas au salarié, la cour d'appel a violé le texte susvisé*".

En l'assumpte Galina Meister⁴⁸³, el TJUE, per una banda, va determinar que la negativa de l'empresa a donar informació -que impedia *per se* que la persona pogués establir un cas *prima facie* de discriminació- podia comptar per l'establiment del cas de *prima facie*. Per altra banda, també va refusar l'existència d'un dret a la informació per part de la persona que presumptivament havia patit discriminació. Això significa que la negativa de donar informació pot comptar com un factor determinant per establir un cas *prima facie*, però també, en canvi, que no està garantida la possibilitat de les víctimes i inclús els òrgans de supervisió per reunir les proves necessàries. De la mateixa manera, en l'assumpte Danfoss, s'estableix que quan una empresa aplica un sistema retributiu caracteritzat per la falta total de transparència, recau sobre la persona empresària la càrrega de provar que la seva política de salaris no es discriminatòria, quan les dones hagin demostrat a partir d'un número relativament important que la seva retribució mitja és inferior a la dels homes⁴⁸⁴. Aquesta jurisprudència aplicada a la discriminació algorítmica suposaria, per exemple, que si un proveïdor no pot o no vol explicar com opera l'algoritme, això pot ser suficient per canviar la càrrega de la prova de la part demandant a la demandada⁴⁸⁵.

La sentència de 6 de febrer de 2020 del Rechtbank Den Haag (Tribunal del Districte de la Haia)⁴⁸⁶, en relació amb el *Systeem Risico Indicatie*⁴⁸⁷ (d'ara en endavant, SyRI) tracta d'un sistema implementat pel Govern dels Països Baixos amb l'objectiu de prevenir i combatre el frau a la seguretat social i a l'agència tributària que consistia en l'elaboració de perfils per a l'assignació individual a cada persona d'un nivell de risc en relació amb la seva probabilitat de cometre frau. L'assignació del nivell de risc es feia en base a la

⁴⁸³ TJUE (Sala Segona), sentència de 19 d'abril de 2012, C-415/10, Galina Meister contra Speech Design Carrier Systems GmbH, par. 46. En aquest cas, la persona demandant al·legava discriminació per raó de sexe, edat i origen ètnic en relació a un procés de selecció de personal i volia obtenir informació sobre quines altres persones candidates eren de l'empresa de selecció.

⁴⁸⁴ TJCE, sentència de 17 d'octubre de 1989, C- 109/88, Handels-og Kontorfunktionærernes Forbund i Danmark contra Dansk Arbejdsgiverforening en nom de Danfoss, par. 16.

⁴⁸⁵ ALLEN, R. i MASTERS, D., "Artificial Intelligence: the right to...", *op. cit.*, 2020, pàg. 597; ALLEN, R. i MASTERS, D., *Regulating for an equal AI...*, *op. cit.*, 2020, pàg. 50. S'estableix en aquest informe que tenint en compte que un sistema d'IA no és suficientment transparent i pot portar a discriminació *prima facie*, les persones que litiguin ho poden fer en base a la jurisprudència del cas Danfoss, on la manca de transparència pot possibilitar la deducció de l'existència de discriminació, essent posteriorment possible que la part demandada es trobi amb grans dificultats per provar justificació objectiva.

⁴⁸⁶ Tribunal del Districte de La Haia, sentència de 5 de febrer de 2020, C/09/550982/ HA ZA 18-388, Case of SyRI. Disponible a: <https://uitspraken.rechtspraak.nl/inziendocument?id=ECLI:NL:RBDHA:2020:1878> Darrera visita el 2 de setembre de 2023.

⁴⁸⁷En català: Sistema Identificador de Riscos.

recopilació i tractament⁴⁸⁸ d'una sèrie de paràmetres i una ingent quantitat de dades personals entre elles el nom, la direcció, el lloc de residència, la direcció postal, la data de naixement, el gènere, les dades laborals, les dades fiscals, les dades comercials, les dades sobre beques rebudes, sobre pensions, sobre l'obligació de reintegrament de les prestacions públiques, dades de l'assegurança de salut, etc.⁴⁸⁹ El Nederlands Juristen Comité Voor de Mensenrechten (Comitè de Drets Humans de Països Baixos) i altres, van plantejar la seva il·legalitat argumentant, entre altres, que es tractava d'un sistema discriminatori enfront persones amb menor renda i col·lectius d'origen immigrant especialment vulnerables⁴⁹⁰. Doncs bé, el tribunal va demanar a la part demandada que aportés informació sobre el model de risc i els indicadors dels quals es composava, com també informació objectivament verificable sobre la naturalesa del sistema algorítmic. La part demandada es va negar a aportar aquesta informació, fet que va ocasionar, entre altres factors, que el Tribunal del Districte de la Haia acabés considerant el sistema contrari a l'article 8.2 del CEDH⁴⁹¹.

Així i tot, és cert que si la part demandant només sospita que la decisió va ser discriminatoria però no pot aportar un element de comparació o prova de tracte discriminatori, la part demandada no té obligació de revelar la lògica subjacent a la decisió. Per tant, si un algoritme genera resultats discriminatoris però les persones afectades no poden provar-ho perquè no tenen accés al resultat d'altres persones, no es tindrà l'obligació de donar accés al sistema ni justificar-lo⁴⁹².

Si la càrrega de la prova canvia, depèn principalment dels tribunals nacionals⁴⁹³. És interessant en aquest sentit la sentència 41/1999, de 22 de març de 1999, on el TC va concedir l'empara a un sindicat perquè el Tribunal Superior de Justícia de Castella i Lleó, en un litigi en el que es discutia la política de contractació gairebé exclusivament en favor dels homes d'una empresa, es va negar a incorporar en el procés certs documents que

⁴⁸⁸La base legal pel tractament provenia de l'article 65.2 de la Llei d'Organització d'Implementació i Estructura d'ingressos (Wet Structuur uitvoeringsorganisatie en inkomen, SUWI) que permet l'elaboració d'informes de riscos per avaluar la possibilitat que la persona física o jurídica faci un ús il·legal de fons governamentals en el camp de la seguretat social i els esquemes relacionats amb els ingressos públics.

⁴⁸⁹EGUÍLUZ CASTAÑEIRA, J.A., "Desafíos y retos que plantean las decisiones automatizadas...", *op. cit.*, 2020, pàg. 351.

⁴⁹⁰Tribunal del Districte de La Haia, sentència de 5 de febrer de 2020, C/09/550982/ HA ZA 18-388, Case of SyRI, FJ 4.8.

⁴⁹¹Tribunal del Districte de La Haia, sentència de 5 de febrer de 2020, C/09/550982/ HA ZA 18-388, Case of SyRI, FJ 6.49.

⁴⁹² SORIANO ARNANZ, A., "Decisiones automatizadas y discriminación...", *op. cit.*, 2021, pàg. 18.

⁴⁹³FARKAS, L. i O'FARREL, O., "Reversing the burden of proof...", *op. cit.*, 2015, pàg. 54.

haguessin acreditat indicis suficients de discriminació (per exemple, l'any en què l'empresa va contractar 120 homes i cap dona), que haguessin permès invertir al càrrega de la prova, corresponent llavors a l'empresa provar que la seva política de contractació era objectiva i aliena a tota discriminació⁴⁹⁴. A partir d'aquesta decisió, el Tribunal Suprem, en sentència de 4 de maig de 2000, va estimar la pretensió del sindicat declarant que l'empresa hauria incorregut en una conducta discriminatòria indirecta en la seva política de contractació, ja que l'empresa no va aportar una justificació objectiva i raonable de per quina raó exigia el títol de Formació Professional II per realitzar unes funcions que no requerien aquesta formació, essent desproporcionat i perjudicial per al col·lectiu femení, que per raons socioeconòmiques no solia posseir aquesta titulació⁴⁹⁵. Interessa aquesta sentència perquè si s'extrapola als sistemes algorítmics evidencia la rellevància del requisit d'aportar indicis sobre la discriminació perquè es dugui a terme la inversió de la càrrega de la prova, i la dificultat que això pot suposar per a les persones que han estat discriminades, que en molts casos necessitaran documentació i informació grupal que els serà inaccessible, ja que no tindran cap tipus de connexió amb la resta de persones que s'han vist discriminades de la mateixa manera, no podent acreditar un nexe d'unió, o que formen part d'un grup desavantatjat en base a una característica comuna.

En el cas de la discriminació mitjançant sistemes algorítmics, donada l'opacitat i l'asimetria d'informació relativa al funcionament del sistema, a una persona li pot ser impossible aportar indicis. En aquest sentit, interessa apuntar que molts estats membres de la UE no disposen d'un procediment d'acció col·lectiva específic pels casos de discriminació⁴⁹⁶, fet que suposa un gran impediment a la protecció efectiva dels drets, donat que aquestes accions donen a les víctimes de discriminació una xarxa i mecanismes de suport que faciliten la realització de tots els tràmits burocràtics i redueixen el cost econòmic de la seva defensa, entre altres aspectes (com veurem en el capítol VII d'aquest treball).

En el següent apartat abordem com la poca probabilitat de determinar la responsabilitat de l'acció discriminatòria dificulta també provar la discriminació algorítmica.

⁴⁹⁴Tribunal Constitucional (Sala Primera), sentència 41/1999, de 22 de marzo de 1999, FJ 9.

⁴⁹⁵Tribunal Supremo (Sala de lo Social, Cuarta), sentència 3663/2000, de 4 de mayo del 2000 (recurso 3708/1999), FJ 6.

⁴⁹⁶CHOPIN, I. i GERMAINE, C., *A comparative analysis of...*, *op. cit.*, 2022, pàg. 95.

4.2. Dificultat d'establir la responsabilitat de l'acció discriminatòria

Tradicionalment la responsabilitat la té la persona o organització que duu a terme l'acció discriminatòria ja sigui directa o indirectament. En el cas dels sistemes algorítmics hi ha una multiplicitat de factors, per exemple, biaixos implícits en les persones que desenvolupen el sistema algorítmic, vetllen pel seu funcionament i apliquen les decisions, bases de dades esbiaixades, una voluntat discriminatòria de l'empresa o organització que empra el sistema, múltiples actors, etc., que poden contribuir a dificultar l'establiment de la responsabilitat.

Si bé és cert que arran d'una llarga construcció jurisprudencial del TJUE en matèria d'igualtat per raó de sexe⁴⁹⁷, s'ha reconegut un nou equilibri probatori que abandona l'element intencional (el TEDH, ja va afirmar en la sentència del cas Nachova i altres contra Bulgària que provar una “motivació racial” podia ser molt difícil a la pràctica⁴⁹⁸), i basa el seu raonament en la constatació objectiva de la diferència de tracte⁴⁹⁹, s'ha de determinar qui respon d'aquesta diferència de tracte, i aquesta determinació, en la discriminació algorítmica, pot ser complicada.

El que ocasiona la dificultat de trobar una persona responsable es pot concretar en què hi ha tantes persones i organitzacions involucrades en la planificació, el desenvolupament i l'ús dels sistemes algorítmics que pot no ser fàcil saber a qui adreçar una demanda davant

⁴⁹⁷ Vegeu: TJCE, sentència de 31 de març de 1981, C-96/80, JP Jenkins contra Kingsgate (Clothing Productions) Ltd.

⁴⁹⁸ TEDH (Gran Sala), sentència de 5 de juliol de 2005, 43577/98 i 43579/98, Nachova i altres contra Bulgària, par. 161.

⁴⁹⁹ AÑÓN ROIG, M. J., “Principio Antidiscriminatorio...”, *op. cit.*, 2013, pàg. 145; RODRÍGUEZ-PIÑERO y BRAVO-FERRER, M., “Los contornos de la discriminación”, *op. cit.*, 2022, pàg. 16. L'autor precisa que la discriminació ja no es limita a la consideració individual de la licitud o arbitrarietat de diferències de tracte per raó de certs atributs protegits, sinó que parteix d'una consideració col·lectiva de fenòmens socials que creen víctimes discriminades pel fet de ser membres d'un determinat grup o categoria de persones. Això implica que la prohibició de discriminació progressivament hagi anat posant l'accent més que en l'element subjectiu de la intenció de qui discrimina, en l'element objectiu de les conseqüències que suposa la diferència adoptada per qui n'és víctima i aquelles persones que es podrien -pel seu atribut protegit- haver trobat en la mateixa situació, de manera que es permet el reconeixement de discriminació inclús en supòsits en els quals falti un element de comparació sobre el qual mesurar una diferència de tracte. Vegeu també les següents sentències del TC, on es constata la irrellevància de l'element intencional, essent suficient fer constar la presència d'un nexa de causalitat adequat entre el comportament antijurídic i el resultat prohibit per la norma: Tribunal Constitucional (Sala Primera), sentència 225/2001, de 26 de noviembre de 2001, FJ 4; Tribunal Constitucional (Sala Segunda), sentència 6/2011, de 14 de febrero de 2011, FJ 2.

un tribunal. Diferents persones o empreses són responsables d'establir els objectius, de construir el procés de decisió, programar i entrenar l'algoritme, recollir i preparar les dades, emprar els algoritmes en la decisió o monitoritzar i supervisar el seu efecte⁵⁰⁰. Es pot detectar el moment en què s'ha produït la discriminació? Actualment no hi ha cap obligació per a una companyia de donar accés als algoritmes o revelar les seves característiques. Aquesta companyia pot haver creat el sistema amb una base de dades accessible públicament, però el valor de les dades -el model creat a partir de la base-, és de propietat privada⁵⁰¹. Els algoritmes d'aprenentatge automàtic o d'aprenentatge profund (i inclús a vegades els que es recolzen en normes que es basen en processos de decisió complexos), són caixes negres i els seus procediments no són intel·ligibles per a les persones, inclús aquelles expertes en aquest àmbit, i es fa difícil traçar i aïllar la causa del resultat discriminatori. Per exemple, les guies de l'Agència Espanyola de Protecció de Dades (d'ara en endavant, AEPD) estableixen que una persona que pren decisions recolzant-se en la intel·ligència artificial no pot evitar la responsabilitat argumentant que hi ha insuficient accés a la informació o coneixement tècnic en relació al funcionament del sistema⁵⁰².

Malgrat això, moltes vegades donada l'opacitat subsisteixen dubtes difícils o impossibles de resoldre per a persones no especialitzades però inclús per a persones especialitzades. Per exemple: quina correlació és problemàtica? El problema ve de les dades esbiaixades o el disseny algorítmic reflexa un estereotip perjudicial? Inclús si es pot localitzar la font del problema, els algoritmes d'aprenentatge automàtic estan en constant evolució i per tant, la font de discriminació pot canviar al llarg del temps⁵⁰³. En aquestes circumstàncies, com es pot concretar la responsabilitat?

La manca de claredat del marc legal, específicament en cas d'aprenentatge automàtic i aprenentatge profund, focalitza les discussions relatives a la responsabilitat a causa de discriminació en sistemes algorítmics en la transparència (imposar una obligació de fer el contingut de l'algoritme accessible), en la rendició de comptes (imposar que aquelles

⁵⁰⁰ COMMITTEE OF MINISTERS OF THE COUNCIL OF EUROPE, *Recommendation CM/Rec(2020)1 of the Committee of Ministers to Member States on the human rights impacts of algorithmic Systems*, Council of Europe, Strasbourg, 8 April 2020, par. 14; ALLEN, R. i MASTERS, D., *Regulating for an equal AI...*, *op. cit.*, 2020, pàg. 70.

⁵⁰¹ CRAWFORD, K., *Atlas of AI*, *op. cit.*, 2021, pàg. 120.

⁵⁰² AEPD, *Adecuación al RGPD de tratamientos que incorporan Inteligencia Artificial. Una introducción*, Madrid, 2020, pàg. 97.

⁵⁰³ GERARDS, J. i XENIDIS, R., *Algorithmic Discrimination in Europe...*, *op. cit.*, 2021, pàg. 74.

persones que planifiquen, desenvolupen, o usen els algoritmes retin comptes) en la interpretabilitat (imposar l'obligació que les persones puguin entendre el fonament d'una decisió algorítmica) o en l'explicabilitat (imposant l'obligació que les persones puguin entendre per què un algoritme produeix un resultat concret i com s'ha obtingut)⁵⁰⁴. Així i tot, quan ens referim a la litigació judicial, no sembla haver-hi unes normes clares i concretes sobre qui ha de considerar-se responsable de l'acció discriminatòria. És essencial que s'asseguri que en tot moment es pugui imputar la responsabilitat a una persona⁵⁰⁵ (tornarem a tractar aquesta qüestió en el l'apartat 4.1 del capítol V des de la perspectiva de l'RGPD i 4.3 del capítol VI des de la perspectiva de la PRIA).

4.3. Dificultat d'establir un cas *prima facie* de discriminació algorítmica d'acord amb els criteris jurisprudencials establerts pel TJUE i el TEDH

En aquest apartat abordem com donat el funcionament de la presa de decisió mitjançant sistemes algorítmics, ens trobem davant la inexistència tant d'un grup desavantatjat, és a dir, moltes persones perjudicades sense eines per interrelacionar-se entre elles (4.3.1), com d'un grup comparador (4.3.2), ambdós necessaris per establir un cas de discriminació indirecta.

4.3.1. Impediments per trobar grup desavantatjat

Com hem explicat en l'apartat 2 del capítol I, els sistemes d'algoritmes classifiquen, prioritzen, associen i filtren conjunts massius de dades per localitzar relacions estadístiques, automatitzant el procés de descobriment de patrons útils que operen regularment de la mateixa manera (model). Aquest model s'aplicarà a casos concrets. Tanmateix, el model es crearà en base a una quantitat massiva de dades de moltes persones, de manera que en el seu funcionament existeix una dimensió col·lectiva no eliminable.

⁵⁰⁴ *Idem*, pàg. 65.

⁵⁰⁵ VILLANI, C., *Donner un sens à l'intelligence artificielle: pour une stratégie nationale et européenne*, Mission confiée par le Premier Ministre Édouard Philippe, Paris, 2018, pàg. 152.

No obstant això, i malgrat que es parteix d'un gran conjunt de dades, el model pot arribar a ser molt precís, de manera que es dona una personalització de tracte inimaginable si no s'empres un sistema algorítmic. Aquesta personalització de tracte serà decisiva en la dificultat de trobar persones que formin part d'un grup desavantatjat. El sistema algorítmic té accés a tantes dades que pot crear respostes individuals per a cada persona que analitza, de manera que serà difícil trobar elements comuns que hagin dut l'algoritme a prendre una decisió discriminatòria.

Tal com la discriminació directa pot provar-se individualment a través de la referència explícita a una característica protegida a la norma impugnada (malgrat moltes altres dificultats, no hi ha necessitat de trobar un grup amb qui comparar-se), la discriminació indirecta requereix comparació grupal⁵⁰⁶. Per consegüent, la problemàtica que expliquem seguidament es dona preferentment en la discriminació indirecta.

En la discriminació indirecta, esdevé molt complicat identificar-se dins d'un grup que s'ha vist desavantatjat, ja que pot ser que no existeixin dos casos en els quals s'hagi actuat de manera idèntica. S'ha de constatar que un grup que comparteix una o diverses característiques s'ha vist en una posició desavantajada en relació a la decisió presa i que les persones d'aquest grup puguin conèixer que existeixen altres en la mateixa situació. Els sistemes de presa de decisió mitjançant algoritmes habitualment discriminaran en base a múltiples factors i serà molt complicat identificar-se-se com a grup desavantatjat concret quan potser els motius de discriminació no són els mateixos per a tots els integrants del grup, convertint en confús definir-se com a grup⁵⁰⁷.

Endemés, normalment el grup desavantatjat està integrat per persones que estan protegides per un motiu de discriminació de manera predominant però no exclusivament⁵⁰⁸. Per exemple, en l'assumpte del TJUE, CHEZ⁵⁰⁹, en un barri poblat majoritàriament per persones d'ètnia gitana (on tenia la demandant una botiga d'alimentació), els cables de distribució d'electricitat estaven situats a una altura de 6 o 7

⁵⁰⁶ WACHTER, S., MITTELSTADT, B. i RUSSELL, C., "Why fairness cannot be automated...", *op. cit.*, 2021, pàg. 16.

⁵⁰⁷ En aquest sentit pot ser rellevant i clarificadora la perspectiva interseccional, que pretén estudiar la discriminació des d'un enfocament sobre com els diferents tipus de discriminació s'afecten entre si, com abordem en el capítol IV d'aquest treball.

⁵⁰⁸ TOBLER, C., *Limits and potential of the concept of indirect discrimination*, *op. cit.*, 2008, pàg. 50.

⁵⁰⁹ TJUE (Gran Sala), sentència de 16 de juliol de 2015, C-83/14, CHEZ Razpredelenie Bulgaria AD contra Komisia za zashtita ot diskriminatsia.

metres perquè ningú s'hi connectés de forma il·legal. La demandant defenia que això li impedia tenir control del comptador del seu consum elèctric i que es tractava d'una discriminació racial –encara que ella no fos d'ètnia gitana. També en l'assumpte Hilde Schönheit⁵¹⁰, el TJUE aborda el càlcul de la pensió de jubilació en base a una tarifa diferent entre les persones treballadores a temps parcial i les persones treballadores a temps complert⁵¹¹. Malgrat aplicar-se aquesta norma a totes les persones treballadores, afectava les dones d'una manera molt més exacerbada, ja que suposaven el 88% de les persones treballadores a temps parcial⁵¹². En resum, tot i estar integrat de manera predominant per dones, el grup desavantajat també comptava amb homes⁵¹³.

La constitució en grup desavantajat no té uns criteris fixes. El TJUE ha suggerit que el grup desavantajat i el comparador han de ser definits en relació a la norma impugnada tal com implementada per la suposada part ofensora⁵¹⁴.

Per mitigar aquesta problemàtica -com també la relativa al grup comparador que abordarem en el següent apartat-, s'ha plantejat emprar la prova estadística, però la jurisprudència del TJUE i el TEDH no sembla adaptada a les necessitats probatòries, com ja hem esmentat en l'apartat 3.3.2 d'aquest mateix capítol, relatiu a la demostració d'un desavantatge particular.

En resum, la discriminació algorítmica dificulta la constitució en grup desavantajat per dur a terme un plet de discriminació indirecta ja que atesa l'opacitat dels sistemes algorítmics i el grau de personalització que poden assolir -malgrat partir de correlacions estadístiques que s'han fet en base a grups de persones-, pot esdevenir complicat definir

⁵¹⁰TJCE (Sala Cinquena), sentència de 23 d'octubre de 2003, C-4/02, Hilde Schönheit contra Stadt Frankfurt am Main.

⁵¹¹Vegeu també: TJUE, sentència de 22 de novembre de 2012, C-385/11, Isabel Elbal Moreno e Instituto Nacional de la Seguridad Social (INSS), Tesorería General de la Seguridad Social (TGSS). El Tribunal estipula que constitueix discriminació indirecta per raó de sexe una legislació estatal que exigeix per aconseguir la pensió de jubilació un període de cotització major a les persones treballadores a temps parcial que a aquelles a temps complert.

⁵¹²Vegeu també, per exemple: TJCE (Sala Sisena), sentència d'11 de setembre de 2003, C-77/02, Erika Steinicke contra Bundesanstalt für Arbeit; TJCE (Sala Primera), sentència de 27 de maig de 2004, C-285/02, Edeltraud Elsner-Lakeberg contra Land Nordrhein-Westfalen.

⁵¹³A escala nacional, vegeu: Tribunal Constitucional (Pleno), sentència 91/2019, de 3 de julio de 2019. El Tribunal va declarar inconstitucional i nul un precepte que regulava el càlcul de la pensió de la jubilació de les persones treballadores a temps parcial per considerar que constituïa una discriminació indirecta per raó de sexe.

⁵¹⁴TJCE, sentència de 13 gener de 2004, C-256/01, Debra Allonby contra Accrington & Rossendale College, Education Lecturing Services, trading as Protocol Professional and Secretary of State for Education and Employment, par. 46.

quin ha estat l'atribut protegit en base al qual el resultat ha estat discriminatori i quines persones conformen el grup desavantajat -sobretot si qui ho ha de descobrir són les persones afectades. Si bé és cert que un grup desavantajat pot veure's reflectit en els resultats de certs sistemes algorítmics (és a dir, per exemple, que com a resultat d'un sistema algorítmic de contractació, es contractin cinquanta homes i una dona), no es pot confiar que els resultats sempre seran visibles, atès que en el cas de ser directament exclòs per part del sistema algorítmic, per exemple, de veure certa publicitat, tothom que formi part del grup desavantajat no en serà conscient (per això, entre d'altres factors, en aquest treball ens decantem per defensar la necessitat d'una protecció col·lectiva-institucional, que arribi allà on les persones individualment considerades no poden arribar, com veurem en el capítol VII i VIII).

4.3.2. Impediments per trobar un grup comparador (absència de *tertium comparationis*)

La construcció d'un estàndard adequat, real o hipotètic, respecte al qual s'estableix el perjudici patit per part de la persona demandant és un dels aspectes més debatuts a l'hora d'analitzar la capacitat del dret antidiscriminatori per donar cabuda a les formes de desigualtat complexes, entre elles la interseccionalitat i la discriminació algorítmica⁵¹⁵.

En el cas del grup comparador, el Tribunal necessita que el terme de comparació sigui apropiat per determinar si el criteri afecta de manera més significativa i perjudicial a un grup -que es pugui determinar en relació a un o diversos trets comuns- que a un altre⁵¹⁶. Per exemple, en el cas *Österreichischer Gewerkschaftsbund*, el TJUE estableix com a no comparable que les dones agafin permisos de maternitat i els homes permisos per a realitzar el servei militar⁵¹⁷. Per establir un grup comparador, la *Méthode Clerc*, emprada per part de la cambra social de la Cour de Cassació francesa⁵¹⁸, sembla ser un sistema efectiu: es defineix una llista de persones treballadores de comparació (que van ser contractades al mateix temps i en les mateixes condicions), a partir de les dades

⁵¹⁵ AÑÓN ROIG, M.J., "Desigualdades algorítmicas...", *op. cit.*, 2022, pàg. 38.

⁵¹⁶ REY MARTÍNEZ, F., *Derecho Antidiscriminatorio*, *op. cit.*, 2019, pàg. 85.

⁵¹⁷ TJCE (Gran Sala), sentència de 8 de juny de 2004, C-220/02, *Österreichischer Gewerkschaftsbund, Gewerkschaft der Privatangestellten contra Wirtschaftskammer Österreich*, par. 64.

⁵¹⁸ TOBLER, C., *Indirect discrimination under Directives 2000/43 and 2000/78*, *op. cit.*, 2022, pàg. 118.

obtingudes, s'elaboren gràfics que mostren els anys i els salaris resultant en corbes de promoció per cada una de les persones de la llista (si no hi ha la informació necessària, la legislació laboral francesa permet sol·licitar la divulgació de les nomines davant del tribunal), i finalment, a partir d'aquesta informació es poden comparar les dades amb la situació de la persona treballadora que presumptivament ha estat discriminada⁵¹⁹.

Per identificar el grup comparador, destaquen dos requeriments pràctics. Primer, amb qui pot/poden comparar-se la/les persona/es demandant/s legítimament, i quina argumentació o evidència es requereix per establir que els dos grups s'haurien de considerar iguals en una mateixa situació. Segon, s'ha d'identificar un grup comparador concret per establir discriminació *prima facie*, o és suficient un comparador hipotètic?⁵²⁰ En relació a la primera qüestió, les comparacions binàries són difícils d'establir. El repte per aquelles persones demandants és que molts atributs considerats tradicionalment com sensibles, ja sigui sexe, ètnia o discapacitat, són construccions socials, la comprensió de les quals és complexa i depèn del context social donat, de manera que comparadors legítims estan necessàriament lligats al context social⁵²¹. Per exemple, com ja hem esmentat, un concepte com “gènere” pot causar l'exclusió de persones que s'identifiquen com a no binàries, transsexuals o intersexuals o forçar-les a enquadrar-se en categories amb els quals no s'identifiquen⁵²².

En relació a la segona qüestió, per identificar grups comparadors apropiats hi ha la possibilitat de plantejar grups comparadors hipotètics. Tanmateix, és problemàtic, com es pot veure en els casos Maniero⁵²³ o , Jyske Finans⁵²⁴, on les persones demandants

⁵¹⁹ *Ídem, ibidem*.

⁵²⁰ WACHTER, S., MITTELSTADT, B. i RUSSELL, C., “Why fairness cannot be automated...”, *op. cit.*, 2021, pàg. 22.

⁵²¹ FARKAS, L. i O'FARREL, O., “Reversing the burden of proof...”, *op. cit.*, 2015, pàg. 41.

⁵²² GERARDS, J. i XENIDIS, R., *Algorithmic Discrimination in Europe...*, *op. cit.*, 2021, pàg. 62.

⁵²³ TJUE (Sala Primera), sentència de 15 de novembre de 2018, C-457/17, Heiko Jonny Maniero contra Studienstiftung des deutschen Volkes Ev. En aquest assumpte, la persona demandant al·legava discriminació indirecta segons la Directiva 2000/43/CE que limitava la possibilitat d'obtenir una beca alemanya a aquelles persones que havien aprovat un examen final exigible a les persones estudiants de dret de la universitat alemanya. La persona demandant, havent cursat el grau en dret en una universitat d'un altre país al·legava que aquest criteri perjudicava a persones en base a l'ètnia. El TJUE va concloure que era necessària una comparació concreta i no una general i abstracta (par. 48).

⁵²⁴ TJUE (Sala Primera), sentència de 6 d'abril de 2017, C-668/15, Jyske Finans A/S contra Ligebehandlingsnavnet. En aquest assumpte, la persona demandant al·legava que la llei danesa, que requeria identificació addicional per la ciutadania de països tercers per obtenir un préstec per comprar un cotxe, causava discriminació indirecta en base a l'origen ètnic. El TJUE va establir que era necessària una comparació específica i concreta i no una comparació abstracta i general (par. 32). També va establir el tribunal, que el país de naixença no justifica, *per se*, la pertinença a un grup ètnic (par. 20) i que no es pot presumir que cada estat té un únic grup ètnic (par. 21).

al·leguen haver patit discriminació indirecta per raó d'ètnia i el TJUE els demana un grup comparador concret i específic en comptes d'un general i abstracte. Tot i així, per posar un exemple -malgrat no sigui del TJUE-, el Tribunal Social Suec va sentenciar un cas relatiu a l'exigència de la llengua sueca per treballar, admetent discriminació indirecta *prima facie* contra qui no tingués la nacionalitat sueca (esdevenint el grup de persones nacionals sueques com a comparador, és a dir, un grup general)⁵²⁵.

A més, identificar grups comparadors apropiats és particularment difícil en casos de discriminació múltiple. Tal com hem esmentat en l'apartat anterior referent a identificar-se com a grup desavantatjat, la dificultat d'establir un grup comparador quan la natura de la discriminació és pluridimensional deixa en una posició molt desemparada a la persona que ha patit discriminació múltiple. S'ha criticat que l'establiment de grup comparador es basa en un únic atribut protegit, ignorant els casos en què, donades les habilitats d'elaboració de perfils dels algoritmes, la discriminació prové de diversos atributs protegits a la vegada⁵²⁶ (com abordem amb més precisió en l'apartat 2.2 del capítol IV). Per tant, s'atomitza tant el grup perjudicat com el grup comparador. Per exemple, és improbable que la discriminació algorítmica es doni únicament per raó de sexe (dones -grup perjudicat- i homes -grup comparador-). Més probablement es donarà combinant atributs, per exemple, per raó de sexe, origen ètnic i edat (dona, afroamericana, de més de 60 anys -grup perjudicat-, i diverses possibilitats de grup comparador).

S'afegeix l'opacitat inherent als sistemes algorítmics⁵²⁷; que els resultats de les persones que es veuen immerses en un tractament de dades estaran protegits per la normativa de protecció de dades⁵²⁸; que l'algoritme, al tenir capacitat d'aprendre, evolucionar i desenvolupar-se de manera autònoma podrà adoptar decisions basades en diferents paràmetres que dificultin l'existència d'un grup comparador; i que, en el cas que el

⁵²⁵FARKAS, L. i O'DEMPSEY, D., *How to present a discrimination claim: handbook on seeking remedies under the EU non-discrimination directives*, European Network of legal experts in the Non-discrimination field, Brussels, 2011, pàg. 37.

⁵²⁶Vegeu: ATREY, S., "Comparators in intersectional discrimination", *Legal Studies*, vol. 38, núm. 3, 2018, pàg. 379-395; GOLDBERG, S., "Discrimination by Comparison", *Yale Law Journal*, vol. 120, 2011, pàg. 728-812.

⁵²⁷La dificultat de comprendre el contingut dels sistemes també impedirà que es pugui considerar discriminatori sense elements de comparació. El TJUE, per exemple, en l'assumpte Dekker, va determinar que rebutjar a una única candidata a un lloc de feina a causa d'estar embarassada era discriminació directa per raó de sexe, ja que només es pot donar el cas si la persona candidata és dona (par. 12). Aquesta situació és difícilment comparable en el cas que la decisió sigui presa per un algoritme al contingut del qual no es pugui accedir.

⁵²⁸HACKER, P., "Teaching fairness to artificial intelligence...", *op. cit.*, 2018, pàg. 1179 i 1182.

sistema no tingui en compte categories sensibles sinó que les infereixi d'altres dades, serà pràcticament impossible deduir per quina raó concreta l'algoritme ha pres la decisió que ha pres, i si un atribut protegit ha tingut alguna incidència en la decisió o no⁵²⁹.

La legitimitat del grup comparador depèn de les circumstàncies del cas, dels arguments presentats i de les resolucions dels tribunals nacionals. Les directives antidiscriminació rarament ofereixen requisits clars que no necessitin futures interpretacions jurisprudencials, fent que la determinació de la legitimitat del grup comparador sigui contextual. De manera que, inclús quan les evidències *prima facie* es poden aconseguir, pot ser un repte pels tribunals avaluar el test de necessitat, proporcionalitat i justificació objectiva que cal fer en la discriminació indirecta⁵³⁰. Per exemple, en un escenari com el d'un procés de selecció, el biaix algorítmic, d'existir, podria acreditar-se per mecanismes probatoris tradicionals (comparació entre individus escollits o exclosos). Ara bé, admès l'indici, i considerant la precisió amb la qual pot operar l'algoritme de selecció, a menys que el biaix es trobi en el propi disseny (codi font), pot ser molt difícil de detectar. Per exemple, en un procés de contractació, on concorren centenars de candidats i s'exclouen grups d'individus diferents amb característiques diverses protegides -o inclús no protegides-, la comparació entre la persona reclamant i les no descartades, no permetrà constatar que la causa de l'exclusió sigui únicament l'atribut protegit o els atributs protegits que posseeix, ja que confluirà tot un univers de característiques inferides que han pogut ser igualment rebutjades⁵³¹. Per tant, la detecció de biaixos discriminatoris obliga a reconsiderar l'estratègia jurídica enfront la tutela de la discriminació producte de mecanismes de decisió automatitzada⁵³².

En aquest punt, el que resulta evident és que malgrat es produeixin discriminacions, només seran aquelles que es puguin traduir en uns resultats concrets del sistema algorítmic, que delimitin clarament un grup perjudicat davant d'un grup (consegüentment) beneficiat, les que seran susceptibles de ser discutides en un litigi, malgrat existeixin moltes altres discriminacions que no es puguin traduir en resultats

⁵²⁹SORIANO ARNANZ, A., "Decisiones automatizadas y discriminación...", *op. cit.*, 2021, pàg. 19.

⁵³⁰ En aquest sentit, vegeu també: TJCE (Sala Segona), sentència de 10 de juliol de 2008, C-54/07, Centrum voor gelijkheid van kansen en voor racismebestrijding contra Firma Feryn NV; TJUE (Sala Tercera), sentència de 25 d'abril de 2013, C-81/12, Asociația Accept contra Consiliul Național pentru Combaterea Discriminării; TJUE (Gran Sala), sentència de 23 d'abril de 2020, C 507/18, NH contra Associazione Avvocatura per i diritti LGBTI - Rete Lenford.

⁵³¹ RIVAS VALLEJO, P., "Sesgos de género en el uso...", *op. cit.*, 2022, pàg. 79.

⁵³² *Ídem*, pàg. 80.

concrets. Anant més enllà, Mulder exposa que el criteri, pràctica o mesura aparentment neutre té un efecte diferent en diferents grups, atès que aquests es troben en situacions diferents, de manera que una comparació no hauria de servir per determinar l'existència de discriminació indirecta⁵³³.

A tall de recapitulació, en aquest apartat hem estudiat les dificultats per aportar prova de discriminació algorítmica. Aquestes dificultats es deuen principalment al caràcter marcadament individual de la protecció antidiscriminatòria que deixa la persona discriminada en situació d'isolació. La manca d'informació i la dificultat per assolir els estàndards d'inversió de la càrrega de la prova, com també la multiplicitat d'actors susceptibles de ser responsables de l'acció discriminatòria, són qüestions problemàtiques en el moment de presentar prova de discriminació algorítmica. També crea dificultats l'establiment tant d'un grup desavantatjat com d'un grup comparador (necessaris en el cas de discriminació indirecta), atesa la capacitat dels sistemes algorítmics d'excloure persones dels resultats directament sense que elles en siguin conscients (com en el cas de la publicitat) i de discriminar -el més probable és que de manera subtil i no evident- en base a múltiples factors que atomitzin els grups perjudicats i comparadors.

5. Conclusions parcials

En aquest capítol ens proposàvem respondre, en part, a la segona qüestió de recerca. Hem abordat per avaluar la suficiència de la normativa antidiscriminatòria, els punts febles que presenta el marc regulador europeu antidiscriminatori per donar eines a les persones que s'han vist discriminades per desafiar la discriminació algorítmica.

Hem trobat diversos punts febles. En primer lloc, si bé es necessari l'ús de categories protegides per instrumentalitzar el dret antidiscriminatori, hi ha categories que poden no ser suficientment representatives o que s'exclouin categories que malgrat no estar protegides per la legislació reforcen desigualtats socials.

⁵³³ MULDER, J., *Indirect sex discrimination in employment. Theoretical analysis and reflections on the CJEU case law and national application of the concept of indirect sex discrimination*, European network of legal experts in gender equality and non-discrimination, Brussels, 2021, pàg. 51.

En segon lloc, les directives antidiscriminació europees també presenten limitacions que suposen la disminució de l'efectivitat de la normativa antidiscriminatòria, ja que només desenvolupen la protecció de sis atributs susceptibles de causar discriminació, i a més, dins aquests atributs, la protecció que s'articula és diferent (per exemple, hi ha limitacions segons l'àmbit material, o possibiliten excepcions d'aplicació en àmbits molt rellevants quan ens referim a les decisions preses mitjançant algoritmes, com la publicitat).

En tercer lloc, la discriminació causada per sistemes algorítmics fa més palès encara el caràcter grupal de la discriminació, en elaborar-se perfils en base a sistemes d'opressió i que aquests a l'aplicar-se a persones concretes produeixin la discriminació. La perspectiva del dret antidiscriminatori anticlassificació, que és la que conforma principalment la normativa, pot no funcionar en els sistemes algorítmics que poden inferir dades sensibles d'altres dades. És a dir, la dissociació de la normativa antidiscriminatòria de la lluita contra les desigualtats sistèmiques que realment operen en la societat, atès que està profundament marcada pel caràcter casuístic i individualitzat de l'aplicació judicial, també suposa un punt feble de la normativa.

En quart lloc, la diferenciació entre discriminació directa i discriminació indirecta establerta a través de la jurisprudència del TJUE sembla no ser operativa quan ens referim a la discriminació algorítmica. Quant a la discriminació directa, atesa l'opacitat inherent a la majoria de sistemes algorítmics, pot ser impossible tant provar que s'ha emprat directament un atribut protegit com que un criteri, malgrat ser formalment neutral, s'ha escollit per raons relacionades amb una categoria protegida per la legislació o que el criteri usat és inseparable o està inextricablement vinculat a una categoria protegida. Quant a la discriminació indirecta, com es basa en provar que un criteri, mesura o pràctica té un resultat discriminatori, sembla més adequada per captar la discriminació algorítmica, però presenta dos problemes: possibilita que es justifiqui objectivament la discriminació i presenta dificultats de prova, sobretot en relació amb l'establiment tant d'un grup desavantajat com d'un grup comparador, ja que els sistemes algorítmics poden excloure persones del resultat directament sense que elles en siguin conscients i discriminar de manera subtil i no evident en base a múltiples factors (atomitzant tant els grups perjudicats com comparadors). Finalment, la normativa és poc clara en relació amb qui és responsable de l'acció discriminatòria quan es produeix mitjançant sistemes algorítmics, i no garanteix a la persona demandant l'accés a la informació per poder invertir la càrrega de la prova.

Per tant, concloem que la normativa antidiscriminatòria europea presenta punts febles que fan que no sigui una regulació que per si sola pugui donar suficients eines a les persones que s'han vist discriminades per desafiar la discriminació ocasionada a través de sistemes algorítmics. Tot seguit abordem la mateixa qüestió però en relació amb la normativa de protecció de dades europea.

Capítol III. Els límits de la regulació de la discriminació mitjançant decisions automatitzades en l'RGPD⁵³⁴

Havent estudiat com i per què les decisions automatitzades mitjançant sistemes algorítmics incrementen les possibilitats de discriminació i els punts febles de la regulació antidiscriminació, ens proposem abordar en aquest capítol, la segona part de la segona qüestió de recerca que ens formulàvem, concretament, si el marc actual de protecció de dades és suficient per fer front a les discriminacions causades per decisions automatitzades.

El dret a la protecció de dades està reconegut a l'article 8 de la CDFUE i l'article 16 del TFUE. També s'inclou en el respecte a la vida privada i familiar reconegut a l'article 7 de la CDFUE i 8 del CEDH. Cal fer menció -donat que també conforma la normativa europea- al Conveni del Consell d'Europa per a la protecció de les persones respecte al tractament automatitzat de dades de caràcter personal (Conveni núm. 108), de 28 de gener de 1981⁵³⁵, modificat el 18 de maig de 2018 (Conveni núm. 108 +)⁵³⁶, que de la mateixa manera que l'RGPD, conté moltes disposicions formulades de manera àmplia que es poden aplicar a diferents situacions i tecnologies⁵³⁷. No obstant això, les relatives a les decisions automatitzades són menys detallades que en l'RDPD⁵³⁸. Per aquesta raó -i

⁵³⁴Aquest Reglament relatiu a la protecció de les persones físiques en allò que respecta al tractament de dades personals i lliure circulació de dades es complementa, a Espanya, amb la Ley Orgánica 3/2018, de 5 de diciembre, de Protección de Datos Personales y garantía de los derechos digitales (LOPDGDD), i tota la normativa sectorial publicada abans i després de l'entrada en vigor de l'RGPD.

⁵³⁵Convenio núm. 108 del Consejo de Europa para la protección de las personas con respecto al tratamiento automatizado de datos de carácter personal, hecho en Estrasburgo el 28 de enero de 1981.

⁵³⁶Convenio núm. 108 del Consejo de Europa para la protección de las personas con respecto al tratamiento automatizado de datos de carácter personal, hecho en Estrasburgo el 28 de enero de 1981, modificado en la sesión núm. 128 del Comité de Ministros en Estrasburgo el 18 de mayo de 2018.

⁵³⁷ZUIDERVEEN, F., *Discrimination, artificial intelligence...*, op. cit., 2018, pàg. 45.

⁵³⁸Transcrivim aquí l'article 9 del Conveni núm. 108 + (la traducció és nostra): "*Drets del titular de dades* 1. Cada individu tindrà dret a: a) no ser subjecte a una decisió que li afecti significativament, basant-se únicament en un tractament automatitzat de dades sense considerar les seves opinions; b) obtenir, quan així ho sol·liciti, en intervals raonables i sense demora o despeses excessives, la confirmació del tractament de les dades personals relacionades amb la seva persona, la comunicació en forma intel·ligible de les dades tractades, tota la informació disponible sobre el seu origen, el període de conservació així com qualsevol altra informació que el responsable del tractament hagi de proporcionar per tal d'assegurar la transparència del tractament conforme a l'article 8, paràgraf 1; c) obtenir, quan així ho sol·liciti, quin és el raonament subjacent al tractament de dades quan els resultats de dit tractament se li apliquin; d) oposar-se en qualsevol moment, en base a la seva situació, al tractament de dades personals que l'involucrin, excepte si el responsable del tractament demostra fonaments legítims per al tractament superiors als seus interessos, drets o llibertats fonamentals; e) obtenir, quan així ho sol·liciti, exempt de costos i sense demores excessives, la rectificació o eliminació, segons el cas, d'aquestes dades si aquests van estar o han estat tractades de forma contrària a les disposicions del present Conveni; f) obtenir una solució jurídica segons

perquè en comptes de ser d'obligada aplicació com l'RGPD, pot ser firmat o no pels estats del Consell d'Europa i també el poden signar altres estats (l'ha firmat, per exemple, l'Uruguai)-, ens centrarem en l'RGPD malgrat fer-hi referència en alguna ocasió.

En relació amb el dret derivat de la Unió Europea, també s'ha de fer menció obligada a la Directiva (UE) 2016/680 del Parlamento Europeu y del Consejo, de 27 de abril de 2016, relativa a la protección de las personas físicas en lo que respecta al tratamiento de datos personales por parte de las autoridades competentes para fines de prevención, investigación, detección o enjuiciamiento de infracciones penales o de ejecución de sanciones penales, y a la libre circulación de dichos datos y por la que se deroga la Decisión Marco 2008/977/JAI del Consejo⁵³⁹ (d'ara en endavant, Directiva 2016/680), transposada a l'Estat Espanyol per la Ley Orgánica 7/2021, de 26 de mayo, de protección de datos personales tratados para fines de prevención, detección, investigación y enjuiciamiento de infracciones penales y de ejecución de sanciones penales⁵⁴⁰ (d'ara en endavant, Llei Orgànica 7/2021). Farem menció puntual tant a la Directiva 2016/680 com a la Llei Orgànica 7/2021 al llarg del treball però la nostra anàlisi se centrarà en l'RGPD no només perquè té un àmbit d'aplicació material més ampli, sinó també perquè l'instrument, un reglament, contràriament a la directiva, és obligatori en tots els seus elements, d'aplicació directa i d'abast general.

En l'RGPD, es fa referència a la discriminació en l'article 22.4 en relació amb l'article 9⁵⁴¹, on s'exigeix una protecció més exhaustiva si el tractament inclou categories

l'article 12 quan els seus drets de conformitat amb el present Conveni hagin estat violats; g) beneficiar-se, sigui quina sigui la nacionalitat o residència, de l'assistència d'una autoritat de control segons el que disposa l'article 15, per a exercir els seus drets de conformitat amb aquest Conveni. 2. El paràgraf 1 no és aplicable si la decisió ha estat autoritzada per una llei a la qual el responsable del tractament està subjecte i s'han establert mesures apropiades per garantir els drets, les llibertats i interessos legítims del titular de dades."

⁵³⁹ Directiva (UE) 2016/680 del Parlamento Europeo y del Consejo, de 27 de abril de 2016, relativa a la protección de las personas físicas en lo que respecta al tratamiento de datos personales por parte de las autoridades competentes para fines de prevención, investigación, detección o enjuiciamiento de infracciones penales o de ejecución de sanciones penales, y a la libre circulación de dichos datos y por la que se deroga la Decisión Marco 2008/977/JAI del Consejo. DOUE, L 119/89, 4.5.2016.

⁵⁴⁰ Ley Orgánica 7/2021, de 26 de mayo, de protección de datos personales tratados para fines de prevención, detección, investigación y enjuiciamiento de infracciones penales y de ejecución de sanciones penales. BOE, núm. 126, de 27 de mayo de 2021. España.

⁵⁴¹ L'article 22.4 de l'RGPD estableix que les decisions automatitzades a les quals es refereix l'apartat 2 de l'article 22 (celebració d'un contracte, decisió autoritzada pel dret de la UE o basada en el consentiment explícit de l'interessat) no es basaran el les categories especials de dades contemplades en l'article 9.1 (aquestes categories són les següents: origen ètnic o racial, opinions polítiques, conviccions religioses o filosòfiques, afiliació sindical, i el tractament de dades genètiques, dades biomètriques dirigides a identificar de manera unívoca a una persona física, dades relatives a la salut o dades relatives a la vida sexual o les orientacions sexuals d'una persona física), llevat que s'apliqui l'article 9.2.a (quan la persona

sensibles de dades; en el considerant 71⁵⁴², on s’aconsella que a fi de garantir un tractament lleial⁵⁴³ i transparent s’apliquin mesures tècniques per corregir inexactituds i errors⁵⁴⁴ que puguin impedir, entre d’altres, efectes discriminatoris; en el considerant 75⁵⁴⁵, on s’aborden els riscos del tractament de dades pels drets i llibertats de les persones físiques; i en el considerant 85⁵⁴⁶, on s’aborden els danys i perjudicis que poden ocasionar les violacions de la seguretat de les dades personals.

interessada hagi donat el seu consentiment específic per una o diverses finalitats especificades i el dret de la Unió Europea o de l’estat membre no estableixi que no pot fer-ho), i lletra 9.2.g (quan el tractament sigui necessari per raons d’interès públic essencial); i s’hagin pres les mesures adequades per salvaguardar els drets i llibertats i interessos legítims de la persona interessada.

⁵⁴² El considerant 71, en la seva literalitat, exposa el següent (la traducció és nostra): “A fi de garantir un tractament lleial i transparent (...), la persona responsable de tractament ha d’emprar processos matemàtics o estadístics adequats per a l’elaboració de perfils, aplicar mesures tècniques i organitzatives adequades per garantir, en particular, que es corregeixen els factors que incorporen inexactituds en les dades personals i es redueix al màxim el risc d’error; assegurar les dades personals de forma que es tinguin en compte els possibles riscos pels interessos i drets de la persona interessada i s’impedeixin, entre altres coses, efectes discriminatoris en les persones físiques (...)”.

⁵⁴³ Cal fer esment al principi de lleialtat, que exigeix fer un ús de les dades que sigui previsible per part de les persones (en relació amb la finalitat buscada) i que no tingui en elles conseqüències adverses injustificades, per tant, existeix un punt de fricció entre els algoritmes de decisió automatitzada i aquest principi atès el potencial discriminatori que poden patir les persones com a conseqüència d’un algorisme que pren decisions esbiaixades. En aquest sentit, vegeu: AUTORITAT CATALANA DE PROTECCIÓ DE DADES, *Intel·ligència artificial. Decisions Automatitzades a Catalunya*, Barcelona, 2020, pàg. 126 i 127.

⁵⁴⁴ Per mitigar la discriminació mitjançant eines tècniques hi ha un gran nombre de mètodes i mètriques existents, vegeu, per exemple: ZEHLIKE, M., HACKER, P., i WIEDEMANN, E., “Matching Code and Law: Achieving algorithmic fairness with optimal transport”, *Data Mining and Knowledge Discovery*, vol. 34, núm. 1, 2020, 163–200; BIRD, S. KENTHAPADI, K., KICIMAN, E. i MITCHELL, M., “Fairness-Aware Machine Learning: Practical Challenges and Lessons Learned”, *Proceedings of the Twelfth ACM International Conference on web search and data mining*, 2019, pàg. 834-835; CASTANEDA, J., JOVER, A., CALVET, L., YANES, S., JUAN, A.A. i SAINZ, M., “Dealing with Gender Bias Issues in Data-Algorithmic Processes: A Social-Statistical Perspective”, *Algorithms*, vol. 15, núm. 9, 2022, pàg. 14 i 15. També, per una visió crítica de l’abast de les eines tècniques per mitigar els biaixos algorítmics, vegeu: DOBBE, R., DEAN, S., GILBERT, T. K., i KOHLI, N., “A Broader View on Bias in Automated Decision-Making: Reflecting on Epistemology and Dynamics”, 2018, pàg. 1-5. Disponible a: <https://arxiv.org/> Darrera visita el 2 de setembre de 2023.

⁵⁴⁵ El considerant 75, en la seva literalitat, exposa el següent (la traducció és nostra): “Els riscos per als drets i llibertats de les persones físiques, de gravetat i probabilitat variables, poden ser deguts al tractament de dades que provoqui danys i perjudicis físics, materials o immaterials, en particular en els casos en els que el tractament pugui donar lloc a problemes de discriminació (...); en els casos en els quals es privi a la persona interessada dels seus drets i llibertats i se’ls impedeixi exercir el control sobre les seves dades personals; en els casos en els quals les dades personals tractades revelin l’origen ètnic o racial, les opinions polítiques, la religió o creences filosòfiques, la militància en sindicats i el tractament de dades genètiques, dades relatives a la salut o dades sobre la vida sexual, o les condemnes i infraccions penals o mesures de seguretat connexes; en els casos en els quals s’avaluïn aspectes personals, en particular l’anàlisi o la predicció d’aspectes referits al rendiment a la feina, situació econòmica, salut, preferències o interessos personals, fiabilitat o comportament, situació o moviments, amb la finalitat de crear o emprar perfils personals; en els casos en els quals es tractin dades personals de persones vulnerables, en particular nens; o en els casos en els quals el tractament impliqui una gran quantitat de dades personals i afecti a un gran nombre de persones interessades”.

⁵⁴⁶ El primer paràgraf del considerant 85, en la seva literalitat, exposa el següent (la traducció és nostra): “Si no es prenen a temps mesures adequades, les violacions de la seguretat de les dades personals poden suposar danys i perjudicis físics, materials o immaterials per a les persones físiques, com la pèrdua de

En l'ordenament espanyol, el dret a la protecció de dades es troba recollit en l'article 18.4 de la Constitució que disposa que la llei limitarà l'ús de la informàtica per garantir l'honor i la intimitat personal i familiar de la ciutadania i el ple exercici dels seus drets⁵⁴⁷. S'ha apostat per l'adaptació de les regulacions existents de la privacitat o la creació de noves regulacions de privacitat per adreçar els problemes generats per les decisions automatitzades mitjançant sistemes algorítmics, per exemple l'RGPD o la Directiva 2016/680, dibuixant-se un marc legislatiu principalment des d'una perspectiva d'anticlassificació, el nucli de la qual és la prohibició (llevat quan sigui estrictament necessari i amb salvaguardes) del processament de certes categories de dades (article 9 de l'RGPD, art. 10 Directiva 2016/680)⁵⁴⁸.

L'RGPD s'aplica a totes les situacions en les quals hi ha recollida o tractament de dades personals⁵⁴⁹ (excepte aquelles que s'emmarquin en l'àmbit d'aplicació de la Directiva 2016/680). Territorialment s'aplica al tractament de dades personals en el context de les activitats d'un establiment del responsable o encarregat en la UE, independentment que el tractament es dugui o no a terme a la UE. Inclús quan un responsable o encarregat del tractament no estigui establert a la UE, s'aplica si les activitats de tractament estan relacionades amb l'oferta de béns i serveis o la monitorització del comportament de les persones (sempre que es dugui a terme a la UE). També s'aplica si el responsable o encarregat del tractament es troba en un lloc en el qual el dret dels estats membres és d'aplicació en virtut del dret internacional públic⁵⁵⁰.

S'aplica l'RGPD⁵⁵¹ a escenaris d'automatització que es basen en dades sobre persones físiques. La norma possibilita que es pugui exigir que es reveli informació sobre el sistema

control sobre les seves dades personals o restricció dels seus drets, discriminació, (...), o qualsevol altre perjudici econòmic o social significatiu per a la persona física en qüestió."

⁵⁴⁷ Article 18.4 de la Constitució Espanyola. El Tribunal Constitucional a través de la seva jurisprudència ha anat desenvolupant aquest article, i estableix actualment que és "*un derecho fundamental autónomo dirigido a controlar el flujo de informaciones que concierne a cada persona, y como derecho fundamental instrumental ordenado a la protección de otros derechos fundamentales, esto es, un instituto de garantía de los derechos a la intimidad y al honor y del pleno disfrute de los restantes derechos de los ciudadanos*". Vegeu: Tribunal Constitucional (Pleno), sentència núm. 76/2019, de 22 de mayo de 2019, FJ 5.

⁵⁴⁸ SORIANO ARNANZ, A., *Data Protection for the prevention...*, *op. cit.*, 2021, pàg. 104, 105 i 108.

⁵⁴⁹ L'RGPD defineix les dades personals (art. 4.1) d'una manera àmplia: qualsevol informació relativa a una persona natural identificada o identificable.

⁵⁵⁰ Article 3 de l'RGPD.

⁵⁵¹ L'RGPD estableix unes condicions mínimes -podríem afegir formals-, que s'han d'assegurar per acreditar la conformitat del tractament realitzat. Entre elles es poden citar l'existència de base per la legitimació del tractament de dades personals (de l'article 6 a l'article 11 de l'RGPD), l'obligació d'informar a les persones de les quals són les dades i ser transparent (de l'article 12 a l'article 14 de l'RGPD), l'obligació de proporcionar a les persones de les quals són les dades els mecanismes per l'exercici dels seus drets (de l'article 15 a l'article 23 de l'RGPD), l'aplicació del principi de responsabilitat proactiva

de presa de decisió, adjudica drets contra decisions específicament automatitzades, i imposa requisits de disseny i de gestió del risc. En aquest capítol estudiem les limitacions que presenten aquestes disposicions per abordar la discriminació, afegint l'estudi a una gran absència com és la regulació vinculant relativa als biaixos algorítmics i analitzant també com el fet que la protecció que proporciona l'RGPD sigui majoritàriament individual no afavoreix l'obtenció d'empara envers la discriminació algorítmica.

Seguim el següent raonament estructurat en cinc apartats. En primer lloc, estudiant l'article 22 de l'RGPD, que regula principalment les decisions automatitzades (1), en segon lloc, analitzant l'absència de regulació vinculant dels biaixos algorítmics (2), en tercer lloc, les mancances de les exigències d'informació i transparència de l'RGPD (3), en quart lloc, les carències del plantejament d'exigències de gestió del risc (4), i en cinquè lloc, els límits de la perspectiva individual de la protecció de l'RGPD envers la discriminació ocasionada pels sistemes algorítmics (5). Finalment, resumim els elements tractats en aquest capítol en un apartat de conclusions parcials (6).

1. Una protecció dèbil i dispersa: article 22 de l'RGPD

L'RGPD va ser aprovat el 27 d'abril de 2016 pel Parlament i el Consell Europeu, derogant així la Directiva 95/46/CE del Parlamento Europeo y del Consejo, de 24 de octubre de 1995, relativa a la protección de las personas físicas en lo que respecta al tratamiento de datos personales y a la libre circulación de estos datos⁵⁵² (d'ara en endavant, Directiva 95/46/CE), que va ser fins llavors la norma europea que regulava la protecció de dades. Va suposar diversos canvis, entre ells la referència en l'article 22 de l'RGPD⁵⁵³ a l'elaboració de perfils que en l'article homòleg de la directiva (article 15)⁵⁵⁴ no apareixia. El primer apartat de l'article 22 -titulat "*decisions individuals automatitzades, inclosa l'elaboració de perfils*"- disposa que tota persona interessada tindrà dret a no ser objecte

(de l'article 24 a l'article 43) i l'assoliment de les condicions per poder realitzar transferències internacionals de dades (de l'article 44 a l'article 50 de l'RGPD).

⁵⁵² Directiva 95/46/CE del Parlamento Europeo y del Consejo, de 24 de octubre de 1995, relativa a la protección de las personas físicas en lo que respecta al tratamiento de datos personales y a la libre circulación de estos datos. DOCE, L 281/31, 23.11.1995.

⁵⁵³S'anomena a aquest article la disposició Kafkiana, en anglès, kafka provision. Vegeu: ZUIDERVEEN, F., *Discrimination, artificial intelligence...*, op. cit., 2018, pàg. 40.

⁵⁵⁴S'ha subratllat en relació a aquest article que no va tenir incidència.

d'una decisió basada únicament en el tractament⁵⁵⁵ automatitzat, inclosa l'elaboració de perfils, que produeixi efectes jurídics en ella o li afecti significativament de manera similar⁵⁵⁶. En el segon apartat s'estipulen les excepcions a la prohibició general, en el tercer apartat s'exposen els drets d'oposició que pot exercir la persona subjecte a una decisió automatitzada (a la intervenció humana, a expressar el seu punt de vista i a impugnar la decisió), i finalment, en el quart apartat es plantegen les exigències del tractament de categories especials de dades.

Per consegüent, i com a base per estudiar com opera (o no) la protecció envers les discriminacions causades per decisions automatitzades, inicialment hem de definir què s'entén per decisió automatitzada i elaboració de perfils en el context de l'RGPD (1.1), analitzant seguidament com opera la prohibició general establerta en l'article i les excepcions (1.2), també com la indeterminació del concepte "*garanties per a la salvaguarda de drets, llibertats i interessos legítims*" pot dificultar la protecció envers la discriminació (1.3), i finalment, la imprecisió i vaguetat de la protecció que es brinda (1.4).

1.1. Decisions automatitzades i elaboració de perfils

Abans d'abordar les definicions de les decisions automatitzades i de l'elaboració de perfils cal determinar què entenem per dades personals, ja que són els elements sense els quals no podrien funcionar. Les dades personals són qualsevol informació relativa a una persona natural identificada o identificable⁵⁵⁷. S'ha de tenir en compte que s'han reconegut les dades com un recurs essencial en relació al creixement econòmic⁵⁵⁸, i que, en conseqüència, no només és rellevant com funcionen els mètodes que s'utilitzen, sinó

⁵⁵⁵ S'entén "tractament" com definit a l'article 4.2 de l'RGPD com qualsevol operació o conjunt d'operacions realitzades sobre dades personals, ja sigui per procediments automatitzats o no, com la recollida, registre, organització, estructuració, conservació, adaptació o modificació, extracció, consulta, utilització, comunicació per transmissió, difusió o qualsevol altre forma d'habilitació d'accés, confrontació o interconnexió, limitació, supressió o destrucció.

⁵⁵⁶ Article 22 de l'RGPD.

⁵⁵⁷ Article 4.1 de l'RGPD; ARTICLE 29 DATA PROTECTION WORKING PARTY, *Opinion 4/2007 on the concept of personal data*, 01248/07/EN, WP 136, Brussels, 20 June 2007, pàg. 6.

⁵⁵⁸ PURTOVA, N., "Do property rights in personal data...", *op. cit.*, 2017, pàg. 69.

també el contingut que empren per dur a terme la seva tasca⁵⁵⁹. Crawford planteja que termes com “*minería de dades*” o frases com “*les dades són el nou oli*” eren part d’un moviment retòric que va fer canviar la noció de “*dada*” com una cosa personal, íntima i subjecte a propietat i control individual, cap a una cosa inerta i no-humana. Les dades van començar a ser presentades com un recurs per ser consumit, un flux per ser controlat i una inversió que cal aprofitar. També s’han començat a veure com un capital, funcionant l’activitat humana com a eina d’extracció de valor⁵⁶⁰. També és rellevant la distinció entre dades personals i no personals. Zuiderveen precisa que part dels processos algorítmics estan fora de l’abast de la llei de protecció de dades, que només s’aplica quan dades personals són tractades, de manera que no s’aplica a models predictius perquè no tenen relació amb persones identificables. Per exemple, si un model predictiu diu que el 80% de les persones que viuen en el codi postal 08094 paga tard les factures, no són dades personals que es refereixen a una persona concreta. En el moment en què aquest model predictiu s’aplica a una persona, la llei de protecció de dades s’aplica un altre cop⁵⁶¹.

En relació a l’article 22 de l’RGPD, per una banda, es defineix com a decisió automatitzada l’habilitat de prendre decisions a través de mitjans tecnològics sense que cap ésser humà hi estigui involucrat⁵⁶². Poden estar basades en qualsevol tipus de dades, aquelles obtingudes directament pels individus implicats (p. e. emplenar un qüestionari), dades que s’han observat sobre un individu (p. e. dades de localització proporcionades per una aplicació) o dades derivades de perfils de persones que ja han estat creats (p. e. capacitat creditícia)⁵⁶³. Per l’altra banda, l’article 4.4 de l’RGPD defineix l’elaboració de perfils⁵⁶⁴ com tota forma de tractament automatitzat de dades personals consistent en utilitzar aquestes per avaluar determinats aspectes individuals de la persona física, en particular per analitzar o predir aquells relatius al rendiment professional, situació

⁵⁵⁹Per aprofundir en relació al concepte de dada personal i la seva regulació en l’RGPD, vegeu: POLO ROCA, A., “Datos, datos, datos: el dato personal, el dato no personal, el dato personal compuesto, la anonimización, la pertenencia del dato y otras cuestiones sobre datos”, *Estudios de Deusto*, vol. 69, núm. 1, 2021, pàg. 211-240.

⁵⁶⁰ CRAWFORD, K., *Atlas of AI*, *op. cit.*, 2021, pàg. 113.

⁵⁶¹ ZUIDERVEEN, F., *Discrimination, artificial intelligence...*, *op. cit.*, 2018, pàg. 44 i 45.

⁵⁶² ALLEN, R. i MASTERS, D., “Artificial Intelligence: the right to...”, *op. cit.*, 2020, pàg. 586.

⁵⁶³ ARTICLE 29 DATA PROTECTION WORKING PARTY, *Guidelines on Automated individual...*, *op. cit.*, 2018, pàg. 8.

⁵⁶⁴Per aprofundir en les implicacions de l’elaboració de perfils, la seva regulació en l’RGPD i els riscos als drets fonamentals de les persones i a la seva dignitat, vegeu: GARRIGA DOMÍNGUEZ, A., “La elaboración de perfiles y su impacto en los derechos fundamentales. Una primera aproximación a su regulación en el Reglamento General de Protección de Datos de la Unión Europea”, *Derechos y Libertades*, núm. 38, Época II, 2018, pàg. 107-139.

econòmica, salut, preferències personals, interessos, fiabilitat, comportament, ubicació o moviments⁵⁶⁵.

El Grup de Treball de l'Article 29, creat per la Directiva 95/46/CE, desenvolupa el concepte d'elaboració de perfils afegint-ne certes precisions. Estableix que, per tal que existeixi, han de convergir tres elements, en primer lloc, ha de ser una forma automatitzada de tractament de dades, en segon lloc, ha de realitzar-se en relació a dades personals i en tercer lloc, amb l'objectiu d'avaluar aspectes individuals de la persona natural, en particular analitzar-los o fer-ne prediccions. Més enllà d'aquests elements precisa que tot i tractar-se d'una forma automatitzada de tractament no exclou que hi participi en part un ésser humà. Es tracta d'un procediment que pot incloure una sèrie de deduccions estadístiques normalment emprades per fer prediccions sobre persones, utilitzant dades de diferents fonts per arribar a una conclusió d'algun aspecte de la persona basat en les qualitats d'altres que apareixen estadísticament similars⁵⁶⁶. L'elaboració de perfils pot dividir-se en tres etapes diferenciades com són la recollida de dades, l'anàlisi automatitzat per identificar correlacions i l'aplicació de la correlació a una persona per identificar característiques del seu comportament present o futur⁵⁶⁷. Els algoritmes s'usen per processos de decisió que afecten a persones gràcies a l'elaboració de perfils, atès que es creen per classificar les persones i posteriorment fer prediccions sobre el seu comportament futur -ja sigui individual o col·lectiu-, essent llavors la base sobre la qual els algoritmes produiran recomanacions per la presa de decisions. També, en si, són decisions automatitzades des del punt de vista que prenen decisions sobre les categories en les quals les dades de les persones tractades seran classificades i els paràmetres sota els quals es mesuraran i avaluaran⁵⁶⁸.

Les decisions automatitzades es poden prendre conjuntament o independentment de l'elaboració de perfils, i de la mateixa manera es pot fer a la inversa, però no són necessàriament activitats separades. Un procés que s'inicia com una presa de decisió

⁵⁶⁵Article 4.4 de l'RGPD.

⁵⁶⁶ARTICLE 29 DATA PROTECTION WORKING PARTY, *Guidelines on Automated individual...*, op. cit., 2018, pàg. 7.

⁵⁶⁷COUNCIL OF EUROPE, *The protection of individuals with regard to automatic processing of personal data in the context of profiling, Recommendation CM/Rec(2010)13 and explanatory memorandum*, Strasbourg, 23 November 2010, pàg. 5.

⁵⁶⁸ SORIANO ARNANZ, A., *Data Protection for the prevention...*, op. cit., 2021, pàg. 62.

automatitzada pot arribar a ser un procés d'elaboració de perfils, depenent de com s'utilitzin les dades⁵⁶⁹.

Definits els dos conceptes, “decisiones automatitzades” i “elaboració de perfils”, cal analitzar els dos elements essencials quan ens referim a les decisions automatitzades tal com estan regulades en l'article 22 de l'RGPD: que es tracti d'una decisió basada únicament en el tractament automatitzat (1.1.1) i que aquesta causi efectes jurídics o que afectin significativament de manera similar⁵⁷⁰ (1.1.2).

1.1.1. Decisió basada únicament en el tractament automatitzat

Segons el Grup d'Experts d'alt nivell en intel·ligència artificial, es pot assolir la supervisió dels sistemes d'intel·ligència artificial mitjançant tres tipus de mecanismes de governança: l'anomenat “human in the loop”, en el qual la persona pot intervenir en qualsevol decisió i en qualsevol moment del cicle de vida del sistema; l'anomenat “human on the loop”, en el qual la intervenció humana es limita al moment del disseny o bé al control del funcionament del sistema; i l'anomenat “human in command” amb una perspectiva econòmica, social, legal i ètica més àmplia, facultant la persona per decidir quan i com usar un sistema, per exemple decidint no usar-lo, establint nivells de discrecionalitat humana durant l'ús del sistema o assegurant la capacitat d'anular una decisió presa pel sistema⁵⁷¹.

L'article 22 de l'RGPD es refereix només a aquelles decisions basades únicament en el tractament automatitzat. En conseqüència, qualsevol decisió en la qual en algun moment hi hagi intervingut una persona, no entrarà en el marge d'aplicació d'aquest article. Això suposa que el tractament automatitzat es pot dur a terme sense les restriccions, limitacions o garanties de l'article 22 quan no es tracta de l'únic mitjà per prendre una decisió. Ens podem preguntar aquí si es poden utilitzar sistemes per simular que no es tracta d'una decisió únicament automatitzada, i així quedar fora de la prohibició de l'article 22 de

⁵⁶⁹ARTICLE 29 DATA PROTECTION WORKING PARTY, *Guidelines on Automated individual...*, op. cit., 2018, pàg. 8.

⁵⁷⁰El mot “similar” no apareixia en l'article 15 de la Directiva 95/46/CE.

⁵⁷¹GRUPO DE EXPERTOS DE ALTO NIVEL SOBRE INTELIGENCIA ARTIFICIAL, *Directrices éticas para una IA fiable*, op. cit., 2019, pàg. 20.

l'RGPD⁵⁷². El Grup de Treball de l'Article 29 especifica que, per considerar que la decisió no ha estat només basada en el tractament automatitzat, és necessari que una persona amb l'autoritat i la competència per canviar la decisió presa pel sistema de tractament de dades –persona responsable del tractament⁵⁷³–, s'asseguri que tota supervisió de la decisió és significativa i que té en compte totes les dades rellevants, de manera que no és una mera formalitat. També es considera necessari que la persona responsable del tractament identifiqui i registri el grau d'intervenció humana en la decisió i el moment en què s'ha produït⁵⁷⁴.

En referència a la jurisprudència de tribunals d'estats membres de la UE, s'argumenta que no s'entra en l'àmbit d'aplicació de l'article 22 quan el poder de decisió últim el té una persona i existeixen altres criteris rellevants a part del sistema automatitzat de presa de decisió i que s'han tingut en compte per prendre la decisió⁵⁷⁵ (considerant-se així que el sistema automatitzat és simplement una eina de suport a la decisió)⁵⁷⁶. Tot i així, al final el problema acaba essent el mateix: la dificultat que l'eina sigui completament objectiva quan sempre hi ha algú que l'ha creat amb un(s) objectiu(s) concret(s).

Quant a la discriminació, suposa un límit a la seva protecció que només s'apliqui l'article 22 quan es tracta de decisions basades únicament en el tractament automatitzat, ja que n'hi ha moltes que es queden fora del seu àmbit d'acció o poden quedar-ne fora si s'articula una aparença d'intervenció humana. Esdevé conflictiva aquesta restricció en

⁵⁷²CASTETS-RENARD, C., “Accountability of Algorithms in the GDPR and beyond: A European Legal Framework on Automated Decision-Making”, *Fordham Intellectual Property, Media and Entertainment Law Journal*, vol. 30, núm. 1, 2019, pàg. 122.

⁵⁷³ Per un anàlisi d'aquesta figura en la doctrina espanyola, vegeu: DURÁN CARDO, B., “Responsable del tratamiento (Comentario al artículo 4.7 RGPD)”, a TRONCOSO REIGADA, A., (dir.) i VILLALBA CANO, L., (coord.), *Comentario al Reglamento General de Protección de Datos y a la Ley Orgánica de Protección de Datos Personales y Garantía de los Derechos Digitales*, Tomo I, Pamplona, Aranzadi, 2021, pàg. 637-665.

⁵⁷⁴ARTICLE 29 DATA PROTECTION WORKING PARTY, *Guidelines on Automated individual...*, op. cit., 2018, pàg. 21.

⁵⁷⁵En aquest sentit, a Catalunya s'utilitza el programa RISCANVI per atorgar o denegar permisos penitenciaris en base a una valoració predictiva de risc de trencament de condemna que no parteix de certs criteris establerts sinó que s'han deduït per part d'un algoritme en base a experiències anteriors, i poden patir modificacions segons les dades que es subministren. La jurisprudència, com es veu reflectida en els autos de l'Audiència Provincial de Girona de 16 de novembre de 2018 (recurs 890/2018) i 20 de desembre de 2018 (recurs 974/2018), entre altres, considera que la forma més oportuna de valorar RISCANVI no es aquella d'emascarar els resultats oferts per l'algoritme de manera automàtica després d'introduir les dades amb les quals es compona el test, sinó la de donar posteriorment una explicació sobre la fiabilitat de les dades, que es produeix al·ludint als factors estàtics que no varien malgrat el transcurs del temps.

⁵⁷⁶BARROS VALE, S. i ZANFIR-FORTUNA, G., *Automated Decision-Making Under the GDPR: Practical Cases from Courts and Data Protection Authorities*, Future of Privacy Forum, Washington, DC, May 2022, pàg. 32.

relació amb el tractament de dades que inclou categories protegides, ja que en el moment en el qual no es basi únicament en el tractament automatitzat, no s'aplicarà tampoc la regulació reforçada, que preveu l'adopció de mesures necessàries per salvaguardar els drets, llibertats i interessos legítims de la persona interessada⁵⁷⁷. Tot seguit abordem els efectes jurídics o d'afectació significativament similar en la persona arran de la decisió automatitzada.

1.1.2. Efectes jurídics o d'afectació significativament similar

Per tal que sigui aplicable l'article 22 de l'RGPD, les decisions han de ser únicament automatitzades i, a part, cal que produeixin efectes jurídics o d'afectació significativament similar en la persona.

Que una decisió produeixi efectes jurídics significa o bé que té una repercussió en els drets de la persona, com per exemple el dret a votar, d'associació, llibertat d'expressió o de tutela judicial, o bé que afecta a l'estatus legal de la persona, com per exemple, refusar l'entrada de la persona en un país o la concessió de la ciutadania, o bé que afecta als seus drets en un contracte, com per exemple, si aquest es cancel·la. També es protegeix quan malgrat la inexistència d'efectes jurídics, la decisió té una repercussió equivalent o significativament similar en el seu impacte. Perquè sigui així, la decisió ha de tenir potencial per afectar significativament les circumstàncies, el comportament o les decisions de les persones interessades, tenir un impacte prolongat o permanent en la persona o conduir a la seva exclusió o discriminació⁵⁷⁸.

El tractament de dades generalment pot tenir un impacte més exacerbat en certs grups de la societat, especialment minories i persones vulnerables. Per exemple si alguna persona amb dificultats econòmiques o que tendeix a tenir-ne és objectiu de publicitat de crèdits a interès alt i al contractar-los incórrer després en deute, o si un banc redueix la línia de crèdit a una persona, no pel seu historial de pagament sinó basant-se en la informació d'altres clients que viuen en la mateixa àrea i compren en els mateixos indrets. En aquest

⁵⁷⁷ Article 22.4 de l'RGPD.

⁵⁷⁸ ARTICLE 29 DATA PROTECTION WORKING PARTY, *Guidelines on Automated individual...*, op. cit., 2018, pàg. 21.

cas s'estaria negant l'accés a oportunitats a causa de dades pertanyents a altres persones⁵⁷⁹ (vegeu apartat 2.1.5 del capítol I sobre la discriminació per associació), i l'efecte seria significatiu.

Existeix la possibilitat d'ocasionar un impacte discriminatori sense produir efectes legals o significatius en la persona subjecte a una decisió automatitzada, és a dir, efectes invisibles o impossibles de concretar-ne la causa, però no per això inexistents. És inevitable preguntar-nos fins a quin punt es pot dibuixar una línia entre el que es considera "significatiu" i el que no. El Grup de Treball de l'Article 29 determina que és difícil ser precís sobre què es podria considerar suficientment "significatiu" com per ser considerat rellevant en referència a l'article 22 de l'RGPD⁵⁸⁰. En relació als exemples que es plantegen per evidenciar un efecte significatiu, Hernández estableix que la determinació ha de realitzar-se de manera casuística, ja que en cada situació han de valorar-se les conseqüències pràctiques, com també, proposa, hauria de considerar-se si la decisió afecta a una persona que ostenta un atribut protegit, ja que una decisió que afecti a una persona que ostenta un atribut protegit pot tenir un efecte significativament major que si no ostenta l'atribut protegit⁵⁸¹. L'ICO⁵⁸² suggereix que "efecte d'afectació significativament similar" es refereix a alguna conseqüència que és més que trivial i potencialment té un resultat desfavorable. També considera que únicament hauria de cobrir aquelles decisions automatitzades en les quals una persona no exerceix influència real en el resultat de la decisió, per exemple quan el resultat de l'elaboració d'un perfil o procés no està avaluat per una persona abans de ser formalitzat com una decisió⁵⁸³.

Per exemple, l'apreciació que fa el tribunal en l'assumpte SyRI sobre els efectes significatius pot evidenciar el caràcter indeterminat del concepte. El tribunal considera que el fet que un informe de risc no necessàriament porti a posteriors investigacions, ni a una sanció administrativa o penal, i que no pugui ser utilitzat com a únic fonament jurídic

⁵⁷⁹*Ídem*, pàg. 22.

⁵⁸⁰*Ídem, ibidem*.

⁵⁸¹HERNÁNDEZ, J.C., "Decisiones algorítmicas de perfilado: Régimen y garantías jurídicas", *Revista Española de Derecho Administrativo*, núm. 203, 2020, pàg. 293-295.

⁵⁸²Sigles en anglès de "Information Commissioner's Office".

⁵⁸³UNITED KINGDOM INFORMATION COMMISSIONER'S OFFICE, *Feedback request profiling and decision making*, London, 6 April 2017, pàg. 19.

d'una decisió dirigida a fer complir una norma, no significa que no produeixi efectes significatius en la vida privada de la persona titular de les dades⁵⁸⁴.

Que la decisió individual sigui únicament automatitzada i tingui efectes jurídics o d'afectació significativament similar no suposa que estigui directament prohibida ja que hi ha possibilitat que s'apliquin les excepcions que s'exposen en l'article 22 de l'RGPD. Interessa abordar-les per tenir una concepció més completa tant de l'estructura de l'article 22 com de les raons que es poden esgrimir per justificar que s'empri un sistema de presa de decisió completament automatitzat.

1.2. Excepcions a la prohibició general

Tal com hem vist, en primer apartat de l'article 22, s'estableix el dret de tota persona a no ser objecte d'una decisió basada únicament en el tractament automatitzat, inclosa l'elaboració de perfils, que produeixi efectes jurídics en ella o li afecti significativament de manera similar. En el segon apartat s'estipulen les excepcions a la prohibició general⁵⁸⁵.

Les excepcions contemplades en el segon apartat impliquen que no s'aplicarà el primer apartat si la decisió presa és necessària per la celebració o execució d'un contracte entre la persona interessada i la responsable del tractament de dades⁵⁸⁶, si la decisió està autoritzada pel dret de la UE o dels estats membres, que s'apliqui al responsable de tractament i que estableixi així mateix mesures adequades per salvaguardar els drets, llibertats i interessos legítims de la persona interessada, o si es basa en el consentiment explícit de la persona interessada⁵⁸⁷. Quan ens referim a la celebració o l'execució d'un contracte entre la persona interessada i la persona responsable del tractament de dades,

⁵⁸⁴Tribunal del Districte de La Haia, sentència de 5 de febrer de 2020, C/09/550982/ HA ZA 18-388, Case of SyRI, FJ 6.59.

⁵⁸⁵Article 22 de l'RGPD.

⁵⁸⁶S'entén "responsable de tractament" tal com definit a l'article 4.7 de l'RGPD on s'estableix que es tracta de la persona física o jurídica, autoritat pública, servei o altre organisme que, sol o conjuntament amb altres, determini les finalitats i mitjans de tractament, especificant que en el cas que sigui en dret de la Unió o dels estats membres qui determini les finalitats o mitjans de tractament també podrà establir la persona responsable de tractament o els criteris específics pel seu nomenament.

⁵⁸⁷ARTICLE 29 DATA PROTECTION WORKING PARTY, *Guidelines on Automated individual...*, op. cit., 2018, pàg. 21.

aquesta última ha d'assegurar-se que el tractament de les dades és apropiat, però a més a més efectiu i el menys intrusiu possible en relació a l'objectiu perseguit. Per tant, davant de l'existència d'altres mètodes, sempre s'ha d'adoptar aquell menys intrusiu en relació a la privacitat⁵⁸⁸. L'apartat b) del segon paràgraf de l'article 22 de l'RGPD – que planteja l'excepció a la prohibició si es tracta d'una decisió autoritzada pel dret de la UE o dels estats membres- obre la possibilitat d'aplicar una gran quantitat de processos automatitzats amb diferents estàndards, ja que dona cert grau de discrecionalitat a cada estat membre, minvant així el potencial harmonitzador de la regulació i les perspectives d'un únic mercat digital⁵⁸⁹ i ocasionant, també, diversitat de regulacions aplicables, en conseqüència, una disminució de la seguretat jurídica. S'afegeix que l'eficiència de l'RGPD depèn de la capacitat dels estats membres de crear els processos i les eines per fer complir la llei⁵⁹⁰, situació que suposa encara més divergències quant a la protecció contra la discriminació que s'atorga a les persones i que evidencia, també, l'existència de límits territorials inexistents en l'àmbit digital essencialment global.

Quan s'empra la decisió únicament automatitzada per l'execució d'un contracte o en base al consentiment explícit de la persona interessada, s'han d'establir mesures adequades per salvaguardar els seus drets, llibertats i interessos legítims. També, per una banda, s'han d'establir mesures de salvaguarda apropiades (com abordem seguidament) quan es tracti d'una decisió autoritzada pel dret de la UE o els estats membres (art. 22.2.b) i per l'altra banda, quan la decisió presa es basi en categories especials de dades establertes en el primer apartat de l'article 9 (origen ètnic, racial, opinions polítiques, etc.) i la persona interessada hagi donat el seu consentiment explícit per al tractament (9.2.a) o sigui necessari per una raó d'interès públic essencial (9.2.g).

1.3. Establiment de mesures adequades per salvaguardar els drets, llibertats i interessos legítims de les persones interessades

⁵⁸⁸BUTTARELLI, G., *Assessing the necessity of measures that limit the fundamental right to the protection of personal data. A Toolkit*, European Data Protection Supervisor, Brussels, 11 April 2017, pàg. 17.

⁵⁸⁹CASTETS-RENARD, C., "Accountability of Algorithms in the GDPR...", *op. cit.*, 2019, pàg. 126.

⁵⁹⁰*Ídem*, pàg. 125-126.

L'establiment de mesures adequades per salvaguardar els drets, llibertats i interessos legítims de les persones interessades comporta, com a mínim, la possibilitat per la persona interessada d'obtenir intervenció humana per part de la persona responsable, d'expressar el seu punt de vista i d'impugnar la decisió, com s'enumera en l'apartat tercer de l'article 22⁵⁹¹. El considerant 71 disposa que les salvaguardes apropiades haurien d'incloure també informació específica de la persona les dades de la qual es tracten per obtenir una explicació de la decisió, i poder impugnar-la⁵⁹². També pot integrar avaluacions freqüents en les bases de dades per buscar biaixos, desenvolupar maneres d'encarar elements perjudicials o realitzar auditories de manera cíclica (no només en el moment del disseny), per introduir procediments i mesures apropiades per prevenir errors, imprecisions o discriminacions en base a categories de dades⁵⁹³.

Un exemple de la discussió sobre l'establiment de les salvaguardes dels drets i llibertats el trobem en la decisió 2018-765DC, de 12 de juny de 2018, del Consell Constitucional francès, on es discuteix, entre altres assumptes, sobre si el poder legislatiu ha establert les garanties apropiades per salvaguardar els drets i les llibertats de les persones sotmeses a decisions administratives individuals preses exclusivament en base a un algoritme⁵⁹⁴. S'al·lega per impugnar l'article que autoritza l'administració a prendre decisions individuals administratives basant-se en un algoritme, que això suposa la renúncia de l'administració a l'exercici de la seva capacitat d'apreciació de les situacions individuals, capacitat que també s'ignoraria en el cas d'algoritmes d'autoaprenentatge capaços de revisar les regles que autoritzen ells mateixos, impedit així a l'administració conèixer les regles en base a les quals s'ha pres la decisió⁵⁹⁵. S'argumenta que tal com està establert pel poder legislatiu, manquen garanties suficients i s'atempta contra els principis de valor constitucional que regeixen l'exercici del poder reglamentari, en la mesura en què no es podrà garantir que les regles establertes per l'algoritme siguin ajustades al dret, afegint-se que atès que les regles les estableix el propi algoritme, tampoc s'estarà complint el principi de publicitat dels reglaments⁵⁹⁶. El tribunal argumenta que s'actua en base a les

⁵⁹¹Article 22.3 de l'RGPD.

⁵⁹²Considerant 71 de l'RGPD.

⁵⁹³ARTICLE 29 DATA PROTECTION WORKING PARTY, *Guidelines on Automated individual...*, op. cit., 2018, pàg. 16; MARTÍNEZ MARTÍNEZ, R. "Inteligencia artificial desde el diseño. Retos y estrategias para el cumplimiento normativo", *Revista Catalana de Dret Públic*, núm. 58, 2019, pàg. 78.

⁵⁹⁴Conseil Constitutionnel, décision 2018-765DC, du 12 juin 2018, considerant 65 a 72. Disponible a: <https://www.conseil-constitutionnel.fr/decision/2018/2018765DC.htm>. Darrera visita el 2 de setembre de 2023.

⁵⁹⁵ *Ídem*, considerant 66.

⁵⁹⁶ *Ídem*, *ibidem*.

regles i criteris que la persona responsable de tractament ha establert, de manera que no s'adopten decisions sense fonament jurídic, i per consegüent, considera definides les garanties apropiades⁵⁹⁷.

En el cas SyRI, el Tribunal del Districte de la Haia es basa en el criteri del TEDH per resoldre fins a quin punt les garanties legals són suficients, establint que depèn de les circumstàncies concretes i de l'efecte de totes les garanties legals combinades, la mesura i el detall de les quals depèn de la gravetat de la interferència. Que la interferència sigui conforme a la llei, pot estar estretament relacionat, al seu torn, amb si és necessària i proporcionada en una societat democràtica⁵⁹⁸. Malgrat seguir essent completament indeterminats els conceptes que han de definir les garanties apropiades, el Tribunal del Districte de la Haia considera, finalment, que la manca d'informació sobre els indicadors de risc i el model de risc i la inexistència de garanties legals per compensar aquesta situació no proporcionen suficients punts de referència per arribar a la conclusió que mitjançant l'ús de SyRI, la interferència amb el dret al respecte a la vida privada és sempre proporcionada i, per tant, necessària⁵⁹⁹, de manera que no s'estan implantant les mesures adequades per salvaguardar els drets, llibertats i interessos legítims de les persones interessades.

Interessa estudiar seguidament com afecta la vaguetat dels conceptes principals que conformen l'article 22 de l'RGPD en la protecció contra la discriminació.

1.4. Vaguetat de l'article 22 de l'RGPD

Per abordar la imprecisió de l'article 22 de l'RGPD creiem oportú seguir amb l'exemple del cas SyRI, on el Tribunal del Districte de La Haia fonamenta tota la seva argumentació en què el sistema implementat pels Països Baixos vulnera l'article 8 del CEDH⁶⁰⁰, que proclama el dret de tota persona al respecte de la seva vida privada i familiar, del seu

⁵⁹⁷ *Ídem*, considerant 71 i 72.

⁵⁹⁸ TEDH (Gran Sala), sentència de 4 de desembre de 2008, 30562/04 i 30566/04, S. and Marper contra Regne Unit, par. 112.

⁵⁹⁹ Tribunal del Districte de La Haia, sentència de 5 de febrer de 2020, C/09/550982/ HA ZA 18-388, Case of SyRI, FJ 6.95.

⁶⁰⁰ Article 8 del CEDH.

domicili i de la seva correspondència⁶⁰¹. Un dels aspectes cardinals de la sentència és el fonament jurídic emprat per establir la il·legalitat del sistema, l'article 8 del CEDH i no la llei de protecció de dades nacional. Inclús el tribunal especifica que no donarà cap opinió sobre el compliment de la definició exacta de la presa de decisions individuals automatitzades a l'RGPD, ni si s'han donat una o més excepcions establertes a l'article⁶⁰². Eguíluz Castañeira entén que hagués estat lògic fonamentar l'argumentació en l'article 22 de l'RGPD; malgrat això, com la seva interpretació segueix essent una incògnita, potser sigui una forma d'evidenciar el dèficits i carències de l'RGPD⁶⁰³.

Reunim seguidament els elements que creiem que poden haver determinat la dificultat per aplicar l'article 22 de l'RGPD. En primer lloc, malgrat considerar-se ja un debat superat en base a les directrius del Grup de Treball de l'Article 29, on es clarifica que l'article 22 estableix una prohibició general per a la presa de decisions basada únicament en el tractament automatitzat, de manera que la prohibició s'aplica independentment que la persona interessada prengui o no una acció respecte del tractament de les seves dades personals⁶⁰⁴, se segueixen donant arguments a favor que és un dret a ser exercit⁶⁰⁵. Aquest dubte de per si, ja ocasiona que les persones que s'han vist discriminades puguin qüestionar-se si és un dret que han de fer valer o hi ha una prohibició que els protegeix *per se*.

En segon lloc, s'ha plantejat que aquest dret pot eludir-se fàcilment en incorporar una intervenció humana no significativa en el bucle algorítmic⁶⁰⁶. També es pot donar el cas que la intervenció humana sigui negligent, més perjudicial o parcial. El Grup de Treball de l'Article 29 clarifica que es requereix una participació que sigui suficientment rellevant com per alterar el curs de la classificació i tingui en compte totes les dades pertinents⁶⁰⁷.

⁶⁰¹Tribunal del Districte de La Haia, sentència de 5 de febrer de 2020, C/09/550982/ HA ZA 18-388, Case of SyRI, FJ 6.7.

⁶⁰²*Ídem*, FJ 6.60.

⁶⁰³EGUÍLUZ CASTAÑEIRA, J.A., “Desafíos y retos que plantean las decisiones automatizadas...”, *op. cit.*, 2020, pàg. 353.

⁶⁰⁴ARTICLE 29 DATA PROTECTION WORKING PARTY, *Guidelines on Automated individual...*, *op. cit.*, 2018, pàg. 19.

⁶⁰⁵TOSONI, L., “The right to object to automated individual decisions: Resolving the ambiguity of Article 22(1) of the General Data Protection Regulation”, *International Data Privacy Law*, vol. 11, núm. 2, 2021, pàg. 161.

⁶⁰⁶WACHTER, S., MITTELSTADT, B. i FLORIDI, L., “Why a Right to Explanation of Automated Decision-Making Does Not Exist in the General Data Protection Regulation”, *International Data Privacy Law*, vol.7, núm. 2, 2017, pàg. 91-93; ZARSKY, T. Z., “Incompatible: The GDPR in the Age of Big Data”, *Seton Hall Law Review*, vol. 47, núm. 4, 2017, pàg. 1016.

⁶⁰⁷ARTICLE 29 DATA PROTECTION WORKING PARTY, *Guidelines on Automated individual...*, *op. cit.*, 2018, pàg. 23.

Aquesta noció és notablement difusa en no especificar el grau d'autoritat de qui determina si la intervenció ha estat significativa. També ens podem plantejar com sabrà la persona que s'ha vist discriminada per una decisió únicament automatitzada que no hi ha hagut intervenció humana. Aquesta qüestió és molt rellevant ja que pot ocasionar que moltes persones no tinguin clar en quin moment i sota quines condicions es pot impugnar la decisió sota l'empareda de l'article 22 de l'RGPD.

En tercer lloc, la discriminació sí que es contempla com efecte d'afectació similar pel Grup de Treball de l'Article 29⁶⁰⁸, però en general es tracta, altre cop, d'una definició vaga, ja que deixa molt marge d'apreciació. Com hem vist, es discuteix en el cas SyRI si es pot considerar que el sistema té efectes d'afectació significativament similar o no.

En quart lloc, entès l'article 22 de l'RGPD com una prohibició i no com a un dret d'oposició, les salvaguardes mínimes s'han d'implementar abans que la decisió automatitzada es prengui. Així i tot, "salvaguardes mínimes" segueix essent un concepte indeterminat donat que es fa en previsió d'ulteriors desenvolupaments nacionals que han de precisar-ne el contingut⁶⁰⁹.

Finalment, en cinquè lloc, si bé s'esmenten el dret a la intervenció humana, a expressar el punt de vista i a impugnar la decisió, és impossible exercir-los sense un grau mínim d'informació que pugui construir un argumentari (s'abordarà aquesta qüestió més específicament en l'apartat 5 d'aquest capítol). A falta d'una posició clara en l'article, s'ha generat un debat sobre si les garanties haurien d'incloure un dret a l'explicació de la presa de decisions algorítmiques, construint una mena de transparència retrospectiva dirigida a la persona afectada, dret que només està anunciat en el considerant 71 de l'RGPD, que no és vinculant⁶¹⁰ (assumpte que tractarem en l'apartat 3 d'aquest capítol). A continuació, estudiem la regulació dels biaixos en l'RGPD.

⁶⁰⁸*Ídem*, pàg. 21.

⁶⁰⁹PARRILLA CASTILLO, J.A., "La Elaboración de Perfiles Políticos en España tras la Sentencia del Tribunal Constitucional de 22 de mayo de 2019 por la que se declara Inconstitucional el Art. 58 bis.1 de la Ley Orgánica del Régimen Electoral General - Perfilado político en España tras el pronunciamiento del Tribunal Constitucional de España de 22 de mayo de 2019", *Anuario do NOVA Consumer Lab*, Año 1, 2019, pàg. 72.

⁶¹⁰HACKER, P. i PASSOTH, J-H, "Varieties of AI Explanations under the Law. From GDPR to the AIA, and beyond", a HOLZINGER, A., GOEBEL, R., FONG, R., MOON, T., MÜLLER, KR. i SAMEK, W. (eds.), *xxAI - Beyond Explainable AI*, Cham, Springer, 2022, pàg. 348.

2. Biaixos algorítmics

Com hem vist, els biaixos algorítmics són un problema cabdal quan ens referim a la discriminació en les decisions automatitzades. Per posar un exemple, en el cas SyRI, el Tribunal considera que l'anàlisi que duu a terme el sistema comporta el risc que es produeixin efectes discriminatoris (involuntaris o no), que es donarien, posem per cas, quan algunes característiques siguin incorrectament atribuïdes a algunes persones (falsos positius), o a la inversa, quan algunes característiques no s'atribueixin a algunes persones quan s'hauria d'haver fet⁶¹¹. També apunta el Tribunal que malgrat acceptar una llista exhaustiva de categories de dades és difícil imaginar qualsevol tipus de dades personals que no siguin elegibles per al tractament a SyRI⁶¹². Estudiem, en aquest apartat, perquè en l'RGPD no es fa esment als biaixos algorítmics en l'articulat (2.1), i com la prohibició de tractament de categories protegides de dades de l'article 9 de l'RGPD, obstaculitza la detecció de biaixos algorítmics (2.2).

2.1. Absència de referència als biaixos en l'articulat

S'aborden els biaixos algorítmics en el considerant 71 de l'RGPD⁶¹³, estipulant que la persona responsable del tractament ha d'emprar processos matemàtics o estadístics adequats per a l'elaboració de perfils, aplicar mesures tècniques i organitzatives adequades que evitin, entre d'altres, els efectes discriminatoris sobre les persones físiques per motius d'origen racial o ètnic, opinió política, religió, conviccions, afiliació sindical, situació genètica o sanitària o orientació sexual, o que donin lloc a mesures que tinguin

⁶¹¹Tribunal del Districte de La Haia, sentència de 5 de febrer de 2020, C/09/550982/ HA ZA 18-388, Case of SyRI, FJ 6.91.

⁶¹²*Ídem*, FJ 6.98.

⁶¹³El considerant 71 disposa que la persona responsable del tractament ha d'utilitzar els procediments matemàtics o estadístics adequats per l'elaboració de perfils, aplicar mesures tècniques i organitzatives apropiades per garantir, en particular, que es corregeixin els factors que introdueixen inexactituds en les dades personals i es redueix al màxim el risc d'error, assegurar les dades personals de forma que es tinguin en compte els possibles riscos pels interessos i els drets de la persona interessada i s'impedeixin, entre altres coses, efectes discriminatoris en les persones físiques per motius de raça o origen ètnic, opinions polítiques, religió o creences, afiliació sindical, condició genètica o estat de salut o orientació sexual, que donin lloc a mesures del mateix efecte.

tal efecte. Com es tracta d'un considerant, no té caràcter vinculant⁶¹⁴. El Grup de Treball de l'Article 29 ha donat poca ajuda pràctica en relació a si el dret a l'explicació pot derivar-se de l'article 22 i el considerant 71 de l'RGPD⁶¹⁵.

D'aquesta manera, únicament s'estableix una línia interpretativa a seguir per part de les persones responsables del tractament però no s'aborda ni s'estableix cap norma imperativa per regular-lo. En referir-nos a la discriminació, aquesta absència suposa que una de les raons principals per la qual es produeix no obtingui l'atenció necessària per part del poder legislatiu. Sembla que la UE, en aquest sentit, s'està inclinant més per una política de la voluntarietat creant uns estàndards d'actuació en base a guies no vinculants, codis de bona conducta i, en fi, normativa facultativa que depèn de la voluntat de les empreses i organitzacions que empen sistemes automatitzats. Podem reconèixer en aquesta línia la guia no vinculant del Grup d'Experts d'Alt nivell en Intel·ligència Artificial, que estableix els principis essencials per desenvolupar la intel·ligència artificial a la Unió Europea⁶¹⁶.

Existeix, en canvi, una prohibició de tractament de categories especialment protegides. S'ha argumentat, com analitzem a continuació, que aquesta prohibició pot dificultar la detecció de biaixos.

2.2. L'RGPD i l'aposta de la protecció en base a categories especials de dades personals

Es precisa en el quart -i últim- apartat de l'article 22, que en relació a les decisions a les quals es refereix el segon apartat, no es basaran en les categories especials de dades personals que es contemplen en l'article 9 de l'RGPD, que prohibeix el tractament de dades que revelin l'origen ètnic o racial, les opinions polítiques, les conviccions religioses

⁶¹⁴WACHTER, S., MITTELSTADT, B. i FLORIDI, L., "Why a Right to Explanation of Automated...", *op. cit.*, 2017, pàg. 80.

⁶¹⁵VEALE, M. i EDWARDS, L., "Clarity, surprises, and further questions in the Article 29 Working Party draft guidance on automated decision-making and profiling", *Computer law and security review*, vol. 34, núm. 2, 2018, pàg. 404.

⁶¹⁶GRUPO DE EXPERTOS DE ALTO NIVEL SOBRE INTELIGENCIA ARTIFICIAL, *Directrices éticas para una IA fiable*, *op. cit.* 2019.

o filosòfiques, o l’afiliació sindical, i el tractament de dades genètiques⁶¹⁷, dades biomètriques⁶¹⁸ dirigides a identificar de manera inequívoca a una persona física, dades relatives a la salut o a la vida sexual o les orientacions sexuals d’una persona física⁶¹⁹ (sense prohibir el tractament de dades que reveli el sexe de la persona). Hi ha certes excepcions, que en relació amb les decisions automatitzades es concreten en autoritzar el tractament de dades quan les persones particulars afectades per aquesta decisió hagin donat el seu consentiment explícit (art. 9.2.a)⁶²⁰, o el tractament sigui necessari per raons d’interès públic essencial, sobre la base del dret de la UE o dels estats membres (art. 9.2.g), que ha de ser proporcional a l’objectiu perseguit, respectar en allò essencial el dret a la protecció de dades i establir mesures adequades i específiques per protegir els interessos i drets fonamentals de la persona interessada⁶²¹.

La regulació espanyola en relació amb l’article 9 no va més enllà que la regulació europea. Com explica Rebollo Delgado, en el tràmit parlamentari de la Ley Orgánica 3/2018, de 5 de diciembre, de Protección de Datos Personales y garantía de los derechos digitales⁶²² (d’ara en endavant, LOPDGDD), es van formular únicament tres esmenes a l’article 9⁶²³

⁶¹⁷S’entén “dades genètiques” segons l’article 4.13 de l’RGPD com dades personals relatives a les característiques genètiques heretades o adquirides d’una persona física que proporcionin una informació única sobre la fisiologia o la salut d’aquesta persona, obtingudes en particular de l’anàlisi d’una mostra biològica de la persona.

⁶¹⁸S’entén “dades biomètriques” segons l’article 4.14 de l’RGPD com dades personals obtingudes a partir d’un tractament tècnic específic, relatives a les característiques físiques, fisiològiques i conductuals d’una persona física que permetin o confirmen la identificació única d’una persona com imatges facials o dades dactiloscòpiques.

⁶¹⁹Sí que serà possible si s’han pres les mesures adequades per salvaguardar drets, llibertats i interessos legítims de la persona interessada i, com s’especifica en el segon apartat d’aquest article novè de l’RGPD en les entrades a) i g) respectivament, la persona interessada ha donat el seu consentiment explícit per al tractament de les seves dades per una o més de les finalitats especificades -excepte quan el dret de la UE o dels estats membres estableixi que la prohibició del primer apartat no pot ser salvada per la persona interessada- o el tractament és necessari per raons d’un interès públic essencial, sobre la base del dret de la UE o dels estats membres, que ha de ser proporcional a l’objectiu perseguit, respectar en allò essencial el dret a la protecció de dades i establir mesures adequades i específiques per protegir els interessos i drets fonamentals de les persones interessades.

⁶²⁰Article 9.2.a) de l’RGPD.

⁶²¹Article 9.2.g) de l’RGPD. El TC basa la inconstitucionalitat de l’article 58 bis.1 de la Ley Orgánica 5/1985, de 19 de junio, del Régimen Electoral General, introduït per la disposició final tercera, segon apartat, de la LOPDGDD, en què no estableix per si mateix les garanties adequades específicament en relació a la recopilació de dades personals relatives a les opinions polítiques pels partits polítics en el marc de les seves activitats electorals, fet que segons el tribunal, constitueix una ingerència en el dret fonamental a la protecció de dades personals de gravetat similar a la que causaria una intromissió directa en el seu contingut nuclear. Vegeu: Tribunal Constitucional (Pleno), sentència 76/2019, de 22 de mayo de 2019, FJ 10.

⁶²²Ley Orgánica 3/2018, de 5 de diciembre, de Protección de Datos Personales y garantía de los derechos digitales. BOE, núm. 294, de 6 de diciembre de 2018. España.

⁶²³Enmiendas núm. 5 a 7, todas realizadas por el Grupo Parlamentario Confederal Unidos Podemos-En común Podem-En Marea. Boletín Oficial de las Cortes Generales, Congreso de los Diputados, Serie A n.º 13-2, 18 de abril de 2018, pàg. 8.

amb escassa fonamentació jurídica, que no va generar debat en profunditat durant la seva tramitació, ni en el Congrés dels Diputats, ni en el Senat⁶²⁴. Per tant, no incorpora un catàleg propi de categories especials de dades i es remet íntegrament a aquelles contemplades a l'RGPD.

L'anàlisi dels efectes de la prohibició de l'ús de certes categories sensibles en el tractament de dades ha revelat que pot, de manera perversa, no només permetre que la discriminació continuï⁶²⁵ sinó participar a que sigui impossible la seva detecció⁶²⁶ (tal com hem esmentat en l'apartat 3.1.1 del capítol I). A més, s'ha plantejat que la perspectiva exhaustiva que adopta l'article 9 de l'RGPD sobre dades sensibles (només es prohibeix l'ús de dades que revelin la informació que enumera l'article 9), deixa fora de l'abast d'aquest article molta informació que també afecta a la vida privada. Per exemple, les dades de geolocalització o els historials de recerca d'internet, poden permetre inferir conclusions molt precises respecte de la vida privada de les persones (hàbits de la vida quotidiana, llocs de residència permanents o temporals, moviments diaris o interessos)⁶²⁷. En aquest cas, com en el dels atributs protegits en el dret antidiscriminatori, sembla que la constitució de categories de dades sensibles -classificacions raonables- és imprescindible per a la instrumentalització del dret a la protecció de dades, malgrat que amb això s'exclouï de la protecció altres categories de dades, l'ús de les quals reforça les desigualtats socials.

S'han proposat solucions o maneres de mitigar el fet que en l'anàlisi massiva de dades totes les dades siguin susceptibles de ser sensibles, com per exemple, Malgieri i Comandé han plantejat la perspectiva "sensitive-by-distance" que pretén detectar quines dades són

⁶²⁴ REBOLLO DELGADO, L., "Las condiciones de tratamiento de categorías especiales de datos (Comentario al artículo 9 RGPD y artículo 9 LOPDGDD)", a TRONCOSO REIGADA, A., (dir.) i VILLALBA CANO, L., (coord.), *Comentario al Reglamento General de Protección de Datos y a la Ley Orgánica de Protección de Datos Personales y Garantía de los Derechos Digitales*, Tomo I, Pamplona, Aranzadi, 2021, pàg. 1037 i 1038.

⁶²⁵ GONZÁLEZ CABAÑAS, J., CUEVAS, A. i CUEVAS, R., "Facebook Use of Sensitive Data for Advertising in Europe", 2018, pàg. 1-15. Disponible a: <https://arxiv.org/>. Darrera visita el 2 de setembre de 2023. S'estima segons aquest estudi, fet arran de la multa d'1,2 milions d'euros de l'AEPD a Facebook, per recollir, emmagatzemar i tractar dades sensibles amb finalitats publicitàries, que malgrat la prohibició establerta en l'article 9 de l'RGPD, a un 73% de les persones usuàries de Facebook a Europa se'ls hi ha assignat preferències publicitàries lligades a categories de dades sensibles.

⁶²⁶ GOODMAN, B. W., "A Step Towards Accountable Algorithms?: Algorithmic Discrimination and the European Union General Data Protection" *29th Conference on Neural Information Processing Systems (NIPS)*, 2016, pàg. 3.

⁶²⁷HUERTA ANGUIANO, J.A., "“Naturaleza Intrínseca”, “Contexto” o “Finalidad” en la...", *op. cit.*, 2020, pàg. 12.

més susceptibles de revelar informació sobre categories sensibles⁶²⁸. També s’han plantejat alternatives per considerar les dades sensibles en base al context en el que es tracten o la finalitat per la qual es tracten. Helen Nissenbaum, entre d’altres, defensa que és el context del tractament el que dota a les dades personals d’una qualitat particular casuística i específica⁶²⁹. L’RGPD, en el considerant 51, fa referència expressament al context del tractament com a causa de diversos riscos pels drets i les llibertats de les persones⁶³⁰. Tanmateix, mai podria ser un criteri únic per regular les categories de dades sensibles justament degut a que la determinació de cert tipus d’informació com a dades personals seria contextual, i no existirien paràmetres que, de manera universal, poguessin ser aplicables per determinar el seu caràcter sensible, augmentant el grau d’incertesa i convertint el concepte en altament subjectiu⁶³¹. Un comitè consultiu del Consell d’Europa ja va plantejar l’any 2004 que l’extremada amplitud del concepte “dades sensibles” feia absolutament necessari abandonar la perspectiva basada en la definició de la naturalesa actual de les dades a favor d’una perspectiva basada en la finalitat del tractament, fent possible considerar el tractament de les dades com a “sensible” en comptes de les dades en sí⁶³².

En conseqüència, les categories de dades que revelen informació personal mai haurien de ser exhaustives, sinó orientatives, deixant la possibilitat d’incorporar o suprimir categories segons l’evolució d’allò que es determina com a sensible comunament per part dels membres d’una societat (existint la possibilitat que hi hagi una falta de correspondència entre les categories previstes en la legislació i allò que les persones titulars de les dades consideren com informació sensible en les seves vides). A més, si bé entenem que l’ordenament jurídic no pot operar en l’àmbit de la subjectivitat (context o

⁶²⁸ MALGIERI, G. i COMANDÉ, G., “Sensitive-by-distance: quasi-health data in the algorithmic era”, *Information and Communications Technology Law*, vol. 26, núm. 3, 2017, pàg. 229-249. Vegeu també en aquest sentit, altres exemples: FELDMAN, M., FRIEDLER, S., MOELLER, J., SCHEIDEGGER, C. i VENKATASUBRAMANIAN, S., “Certifying and Removing Disparate Impact”, *Proceedings of the 21st ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 2015, pàg. 1-28; ŽLIOBAITĖ, I., CUSTS, B., “Using sensitive personal data may be necessary for avoiding discrimination in data-driven decision models”, *Artificial Intelligence Law*, vol. 24, núm. 2, 2016, pàg. 183-201.

⁶²⁹NISSENBAUM, H., *Privacy in Context: Technology, Policy, and the Integrity of Social Life*, Stanford, Stanford University Press, 2009, pàg. 129 i s.

⁶³⁰Considerant 51 de l’RGPD.

⁶³¹HUERTA ANGUIANO, J.A., ““Naturaleza Intrínseca”, “Contexto” o “Finalidad” en la...”, *op. cit.*, 2020, pàg. 18-20.

⁶³²POULLET, Y. i DINANT, J.-M., *Report on the application of data protection principles to the worldwide telecommunication network*, Consultative committee of the Convention for the protection of individuals with regard to automatic processing of personal data, T-PD (2004) 04 final, Strasbourg, 12 December 2004, pàg. 43.

finalitat), és cert que aquestes perspectives subratllen la necessitat de fer més extensa la interpretació del concepte de manera que la protecció doni més garanties⁶³³ en vista d'un futur en el qual es perfila de manera més evident que totes les dades són “sensibles” segons la lectura que se'n faci d'elles⁶³⁴.

Resumidament, que l'RGPD no reguli en l'articulat els biaixos algorítmics és un punt feble per tal d'abordar la discriminació algorítmica. A més, la perspectiva adoptada en l'article 9 de l'RGPD de prohibir en el tractament de dades l'ús de certes dades personals, pot dificultar la detecció de biaixos, atesa la capacitat de l'anàlisi massiu de dades per inferir dades personals d'altres dades. Per això s'han proposat alternatives per captar quines dades, malgrat que no ho semblin, poden ser sensibles, o per establir la sensibilitat de les dades segons el context del qual s'extreuen o la finalitat per la qual s'usen. Tot seguit estudiem la perspectiva des de la qual s'aborda l'accés a informació i la transparència en l'RGPD.

3. Informació i transparència

L'RGPD planteja certes exigències d'informació i transparència. La informació que puguin obtenir les persones que són objecte de decisions automatitzades pot tenir una incidència directa en la detecció de situacions discriminatòries i la possibilitat d'actuar al respecte. Analitzem, en aquest apartat, primerament quines exigències d'informació i transparència estipula l'RGPD (3.1), i segonament, quins obstacles poden trobar les persones que són subjectes de decisions automatitzades mitjançant sistemes algorítmics per obtenir la informació (3.2).

⁶³³HUERTA ANGUIANO, J.A., ““Naturaleza Intrínseca”, “Contexto” o “Finalidad” en la...”, *op. cit.*, 2020, pàg. 27.

⁶³⁴Vegeu, en aquest sentit: PURTOVA, N., “The law of everything: Broad concept of personal data and future of EU data protection law”, *Law, Innovation and Technology*, vol. 10, núm. 1, 2018, pàg. 40–81, esp. 80. L'autora defensa que la distinció entre dades personals i dades no personals s'hauria de suprimir.

3.1. Com es regula l'accés a la informació i la transparència en l'RGPD?

En l'article 5 de l'RGPD s'estableixen els principis relatius al tractament que involucri dades personals, que segons Kaminski són una versió de les Fair Information Practices (FIP)⁶³⁵. En el primer apartat, s'exposa que el tractament s'ha de realitzar d'una manera lícita, lleial i transparent⁶³⁶. També, segons l'article 12.1 de l'RGPD la persona responsable del tractament ha de donar a les persones informació concisa, transparent, intel·ligible i de fàcil accés sobre el tractament de les seves dades; però no concreta tècniques per fer-ho⁶³⁷. L'article 13 i l'article 14 determinen concretament la informació que cal donar a la persona que cedeix les seves dades. Quan es tracta d'informació obtinguda directament de la persona detentora de les dades, els requisits d'informació⁶³⁸ han de ser acomplerts en el moment en el qual s'obtenen les dades⁶³⁹, en canvi, quan es tracta de dades obtingudes indirectament, és a dir, quan no les ha lliurat la persona, l'article 14.3 de l'RGPD estableix uns terminis determinats per fer-ho⁶⁴⁰. L'article 15 estableix el dret de la persona que veu tractades les seves dades a tenir-ne accés i a certa

⁶³⁵KAMINSKI, M.E., "Binary Governance...", *op. cit.*, 2019, pàg. 1585. Les FIP són el conjunt de principis desenvolupats durant els últims cinquanta anys per modelar les normes de bones pràctiques relatives a les dades personals. Els principis fonamentals són de Notificació/Consciència, Elecció/Consentiment, Accés/Participació, Integritat/Seguretat i Execució/Reparació.

⁶³⁶Article 5 de l'RGPD.

⁶³⁷Article 12.1 de l'RGPD.

⁶³⁸L'article 13 de l'RGPD disposa la informació que s'haurà de facilitar quan les dades personals s'obtinguin de la persona interessada, de manera que la persona responsable del tractament haurà de comunicar la identitat i les dades de contacte de la persona responsable, les dades de contacte del delegat de protecció de dades, les finalitats del tractament a les quals es destinaran les dades personals i la base jurídica del tractament, si és el cas, els interessos legítims de la persona responsable o del tercer, els destinataris o les categories de destinataris de les dades personals, i la intenció, si és el cas, de la persona responsable de transferir dades personals a un tercer país o organització internacional. També es proveirà a la persona interessada, en el moment en el qual s'obtinguin les dades personals, el termini durant el qual es conservaran les dades personals, l'existència del dret a sol·licitar accés a les dades, a rectificar-les o a suprimir-les, a limitar o oposar-se al seu tractament, així com el dret a la portabilitat de les dades, també l'existència, si es dona el cas, del dret a retirar el consentiment en qualsevol moment sense que afecti a la licitud del consentiment previ a la retirada, el dret a presentar una reclamació davant l'autoritat de control, si la comunicació de les dades és un requisit legal o contractual, l'existència de decisions automatitzades inclosa l'elaboració de perfils, informació sobre la lògica aplicada i la importància i les conseqüències previstes del tractament, i en el cas de l'existència d'un tractament ulterior, la persona responsable de tractament haurà de proporcionar informació sobre les finalitats i qualsevol informació addicional.

⁶³⁹Article 13 de l'RGPD.

⁶⁴⁰L'article 14.3 de l'RGPD estableix que els terminis per informar a la persona és d'un mes com a màxim tenint en compte les circumstàncies específiques del tractament de dades, si s'han d'utilitzar per comunicar-se amb la persona interessada, en el primer moment en el que es contacta amb ella, i si s'han d'utilitzar per comunicar-se amb un altre destinatari, en el moment en el qual les dades personals siguin comunicades per primera vegada.

informació⁶⁴¹. Tant l'article 13.2 (f), com l'article 14.2 (g) i l'article 15.1 (h) de l'RGPD exigeixen en el cas que el tractament involucri decisions automatitzades informar la persona interessada de les tres qüestions següents: (i) l'existència de decisions automatitzades; (ii) la possibilitat d'obtenir informació significativa sobre la lògica aplicada i; (iii) la possibilitat d'obtenir informació sobre la importància i les conseqüències previstes del tractament a la persona interessada⁶⁴².

La informació que la persona responsable del tractament de les dades ha de proporcionar ha de ser d'una tipologia que conscienciï la persona sobre el tipus de tractament que s'està duent a terme amb les seves dades i li proporcioni certesa i confiança en els resultats⁶⁴³. Com explica Castañer Codina, aquesta informació ha d'incloure la importància i les conseqüències que previsiblement té el tractament per a la persona. Aquests termes s'han interpretat com el deure de facilitar informació sobre el tractament previst i futur, i sobre com pot afectar la decisió automatitzada a la persona interessada. Per exemple, en el cas de valorar-se la seva solvència, el primer que ha de conèixer la persona interessada són els efectes de la seva negativa a sotmetre's a l'avaluació que, com és fàcil de suposar, normalment implicaran que l'entitat financera no li concedeixi el crèdit⁶⁴⁴.

El considerant 71 de l'RGPD és l'única disposició que explícitament menciona la possibilitat de rebre explicació sobre la decisió presa i no revela ni la finalitat d'aquesta ni el seu contingut⁶⁴⁵.

L'eficàcia dels drets que s'atorguen a la persona en l'article 22.3 (d'intervenció humana, d'expressar el seu punt de vista i d'oposició), i l'efectivitat de la protecció envers la discriminació algorítmica, estan estretament lligades a la transparència del funcionament del sistema de presa de decisió i a la informació de rellevància per poder rebatre'l que pot

⁶⁴¹Article 15 de l'RGPD. Per una perspectiva crítica en relació amb l'efectivitat de l'accés a les dades personals, vegeu: HILDEBRANDT, M. "The Dawn of a Critical Transparency Right for the Profiling Era" a BUS, J. *et al.* (eds.), *Digital Enlightenment Yearbook*, Amsterdam, IOS Press BV, 2012, pàg. 49-54. L'autora defensa que no és tan important l'accés a les dades personals com l'accés als perfils (generats en base a dades personals de moltes persones) que s'han inferit a partir d'elles.

⁶⁴²Article 13.2 (f), 14.2 (g) i 15.1 (h) de l'RGPD.

⁶⁴³ AEPD, *Adecuación al RGPD de tratamientos que incorporan Inteligencia Artificial. Una introducción*, Madrid, 2020, pàg. 24.

⁶⁴⁴CASTAÑER CODINA, J., "La evaluación de la solvencia de las personas mediante el uso de algoritmos" a HUERGO LORA, A.J., (dir.) i DÍAZ GONZÁLEZ, G. M., (coord.), *La regulación de los algoritmos*, Pamplona, Aranzadi, 2020, pàg. 270.

⁶⁴⁵ WACHTER, S., MITTELSTADT, B., i RUSSEL, C., "Counterfactual explanations without opening the black box: Automated decisions and the GDPR", *Harvard Journal of Law & Technology*, vol. 31, núm. 2, 2018, pàg. 880.

obtenir la persona que s'hi veu subjecte. Oferir una referència tècnica sobre l'algoritme pot ocasionar l'efecte contrari, essent una informació opaca, confusa i que condueixi a la fatiga informativa. De manera que, veure com funciona internament el sistema no suposa que s'entengui o es controli⁶⁴⁶. Per tant, l'RGPD només és un instrument legal significatiu per donar suport a la igualtat si els nivells adequats de transparència existeixen⁶⁴⁷. Tot seguit abordem els obstacles per assolir una informació eficient sobre les decisions que empren sistemes algorítmics en base a l'articulat de l'RGPD.

3.2. Obstacles per l'obtenció d'informació i transparència a causa de les decisions automatitzades mitjançant sistemes algorítmics

Quan ens referim a la discriminació en les decisions automatitzades, molts factors poden impedir que la informació i la transparència siguin instruments efectius per fer-hi front.

La transparència es pot veure limitada per restriccions que impedeixen la divulgació d'informació i la propietat intel·lectual i industrial. Malgrat que el Grup de Treball de l'Article 29 hagi interpretat que els drets de propietat intel·lectual no poden ser invocats per denegar l'accés o refusar el proveïment d'informació a la persona que cedeix les seves dades⁶⁴⁸, l'RGPD no planteja excepcions o limitacions quant a l'abast dels drets de propietat industrial, ocasionant moltes dificultats envers la transparència del sistema i impossibilitant, en molts casos, que la persona que es veu subjecte a una decisió automatitzada rebi informació significativa sobre la lògica del sistema utilitzat⁶⁴⁹. En el cas SyRI, per exemple, un dels arguments que esgrimeix el Tribunal del Districte de la Haia en relació amb la manca de transparència, és que el sistema en cap cas proporciona informació sobre les dades que poden demostrar la causalitat d'una certa circumstància, és a dir, quines dades objectives poden conduir de manera justificada a la conclusió que

⁶⁴⁶AEPD, *Adecuación al RGPD de tratamientos que incorporan Inteligencia Artificial. Una introducción*, Madrid, 2020, pàg. 24.

⁶⁴⁷ ALLEN, R. i MASTERS, D., *Regulating for an equal AI...*, *op. cit.*, 2020, pàg. 24.

⁶⁴⁸ARTICLE 29 DATA PROTECTION WORKING PARTY, *Guidelines on Automated individual...*, *op. cit.*, 2018, pàg. 17.

⁶⁴⁹Article 15.4 RGPD i Considerant 63 RGPD. Vegeu també: CASTETS-RENARD, C., "Accountability of Algorithms in the GDPR...", *op. cit.*, 2019, pàg. 119.

existeix un major risc de frau⁶⁵⁰. Inclús, el Tribunal exposa que és el propi Estat que deliberadament escull no revelar la informació relativa al model de risc ni als indicadors dels quals està compost o pot estar compost el model de risc, motivant la seva negativa en què, si es divulgués, les persones podrien ajustar la seva conducta per enganyar el sistema⁶⁵¹.

Una altra casuística pròpia de les decisions automatitzades que també pot causar la pràctica ineficàcia dels articles 13.2 (f), 14.2 (g) i 15.1 (h) de l'RGPD, en relació amb informar sobre l'existència d'una decisió automatitzada, és el cas en què la decisió presa mitjançant un sistema algorítmic hagi exclòs la persona (per exemple, de visualitzar un anunci), ja que aquesta persona no sap que està formant part de la presa d'una decisió.

També la complexitat interna dels sistemes pot no ser reduïble, ja sigui per l'escala de l'operació, ja sigui per les dificultats que suposa explicar el coneixement tècnic al públic general⁶⁵², impedint l'efectivitat dels articles 13, 14 i 15 de l'RGPD. També, l'efectivitat es veu compromesa quan no hi ha un llenguatge comú entre qui proporciona la informació i qui la rep (llenguatge algorítmic-llenguatge humà). Ens referim aquí a que malgrat els algoritmes són una simplificació en dades de realitats més complexes, no és fàcil explicar com la interacció d'aquestes dades ha obtingut resultats aplicables altre cop a la realitat complexa. Malgrat que quan ens referim a la intel·ligència artificial, en tant que multitud interrelacionada de sistemes de poder, la transparència completa sigui impossible⁶⁵³, Kaminski argumenta que s'ha de treballar per la seva existència, atès que és necessària per a la justificació d'un sistema: ningú pot desafiar la validesa o legitimitat d'una decisió si no pot veure el raonament en què s'ha basat⁶⁵⁴. Com s'ha subratllat, moltes formes d'opacitat dels sistemes d'aprenentatge automàtic han donat coartada a algoritmes defectuosos⁶⁵⁵ i decisions discriminatòries⁶⁵⁶.

⁶⁵⁰Tribunal del Districte de La Haia, sentència de 5 de febrer de 2020, C/09/550982/ HA ZA 18-388, Case of SyRI, FJ 6.87.

⁶⁵¹*Ídem*, FJ 6.49.

⁶⁵²BURRELL J., "How the machine 'thinks': Understanding opacity...", *op. cit.*, 2016, pàg. 5.

⁶⁵³ CRAWFORD, K., *Atlas of AI, op. cit.*, 2021, pàg. 12.

⁶⁵⁴ KAMINSKI, M.E., "Binary Governance...", *op. cit.*, 2019, pàg. 1546.

⁶⁵⁵BRENNAN-MARQUEZ, K. i HENDERSON, S., "Artificial Intelligence and Role-Reversible Judgment", *Journal of Criminal Law and Criminology*, vol. 109, núm. 2, 2019, pàg. 146.

⁶⁵⁶SÁNCHEZ-MONEDERO, J., DENCİK, L., i EDWARDS, L., "What Does It Mean to "solve" the Problem of Discrimination in Hiring?: Social, Technical and Legal Perspectives from the UK on Automated Hiring Systems", *FAT* '20: Proceedings of the 2020 Conference on Fairness, Accountability, and Transparency*, 2020, pàg. 460.

A més, existeix un debat sobre l'accés al codi font de diferents sistemes automatitzats emprats pel sector públic⁶⁵⁷. S'entén que els estàndards de transparència en el sector privat seran molt més flexibles, tot i no estar obert a debat judicial encara. A França, la Loi n° 2016-1321, du 7 octobre 2016, pour une République numérique⁶⁵⁸, va introduir en el Code des relations entre le public et l'administration⁶⁵⁹, l'article L312-1-3 que estableix l'obligació de publicar en línia per part de certes administracions públiques franceses - llevat dels secrets protegits- les regles que defineixen els principals tractaments algorítmics utilitzats en el desenvolupament de les seves funcions quan siguin la base de decisions sobre les persones⁶⁶⁰. Diversos tribunals i òrgans administratius han obligat a donar accés al codi font de l'algoritme⁶⁶¹, així i tot, en ser pocs, no es pot entendre com una tendència consolidada. En la sentència del Tribunal Administratiu Regional (TAR) de Lazio, de 22 de març de 2017, s'argumenta per reconèixer al demandant el dret d'accedir al programa amb el qual s'havia pres una determinada decisió (aplicació informàtica amb la qual es decidien els trasllats interprovincials del personal docent), que en la mesura en què les prediccions algorítmiques produeixen efectes jurídics per integrar-se en la presa de decisions administratives, és obligatori atorgar accés als algoritmes, ja que de no ser així el control de la decisió serà ineficaç⁶⁶².

⁶⁵⁷ BOIX PALOP, A., “Los algoritmos son reglamentos: la necesidad de extender las garantías propias de las normas reglamentarias a los programas empleados por la administración para la adopción de decisiones”, *Revista de Derecho Público: Teoría y Método*, vol. 1, 2020, pàg. 243. L'autor exposa que en el sector públic, en el qual s'apliquen majors exigències de transparència en les seves accions, ens trobem amb què no es reconeix el dret a accedir al codi font, ni es preveu en cap cas que sigui necessària la seva publicació.

⁶⁵⁸ Loi n° 2016-1321 du 7 octobre 2016 pour une République numérique. JORF n°0235 du 8 octobre 2016. France.

⁶⁵⁹ Ordonnance n° 2015-1341 du 23 octobre 2015 relative aux dispositions législatives du code des relations entre le public et l'administration. JORF n°0248 du 25 octobre 2015. France.

⁶⁶⁰ L'article L312-1-3 del Code des relations entre le public et l'administration disposa literalment que: “*Sous réserve des secrets protégés en application du 2° de l'article L. 311-5, les administrations mentionnées au premier alinéa de l'article L. 300-2, à l'exception des personnes morales dont le nombre d'agents ou de salariés est inférieur à un seuil fixé par décret, publient en ligne les règles définissant les principaux traitements algorithmiques utilisés dans l'accomplissement de leurs missions lorsqu'ils fondent des décisions individuelles*”.

⁶⁶¹ L'obligació de revelar els algoritmes emprats per l'administració pública va ser establerta per la Comissió Catalana de Garantia del dret d'accés a la informació pública en la Resolució de 21 de setembre de 2016, d'estimació de les Reclamacions 123/2016 i 124/2016 (acumulades), ja que es va considerar que la confidencialitat del procés de decisió no podia servir com una base legítima per prevenir revelar un algoritme que seleccionava el professorat que qualificava els exàmens d'accés a la universitat. També, a França, la Commission d'accès aux documents administratifs, en les decisions núm. 20144578 i núm. 20180276 de 8 de gener de 2015 i de 19 d'abril de 2018 respectivament, va garantir l'accés als programes emprats per calcular els diferents aspectes de la situació econòmica de persones i llars, ja que es va considerar que aquests sistemes automàtics constituïen documents administratius.

⁶⁶² Tribunal Regional Administratiu de Lazio-Roma (Secció tercera bis), sentència núm. 3769, de 22 de març de 2017.

A l'Estat Espanyol, Civio ha plantejat judicialment la impugnació de BOSCO (l'algoritme que el Ministeri per a la Transició Ecològica va llançar el 2017 per determinar qui pot percebre ajuda financera en la factura de la llum a través de la revisió de sol·licituds, que ha rebut gran quantitat de denúncies per denegar l'ajuda a persones que complien tots els requisits, així com per falta de transparència), en base a la negativa del Consell de Transparència i Bon Govern de proporcionar l'accés al codi font de l'algoritme. L'Administració justifica amb la propietat intel·lectual la negativa a l'accés, ja que el codi font s'expressa de forma escrita i pot protegir-se com a obra literària⁶⁶³. Civio argumenta, en el seu recurs contenciós-administratiu presentat l'any 2019, que negar l'accés al codi font impedeix que la ciutadania pugui verificar si una eina a través de la qual se li aplica el dret és conforme a la llei o la incompleix, al·legant que qualsevol sistema emprat per l'administració pública ha de fer-se públic per defecte⁶⁶⁴. L'Administració respon, en la seva contestació a la demanda, que l'algoritme es limita a recollir els criteris legals que permeten el reconeixement del dret a l'ajuda pública, però qui té el poder decisorí és l'organisme públic i que la decisió en tot cas és revisable per a la persona interessada seguint els canals habituals que no es veuen obstaculitzats per l'algoritme⁶⁶⁵. El jutjat contenciós-administratiu, en sentència de 30 de desembre de 2021⁶⁶⁶ va resoldre denegant l'accés al codi font de l'algoritme en base a, entre d'altres, els límits del dret d'accés que posa l'article 14.1, lletres d), g), i), j), k) de la Ley 19/2013, de 9 de diciembre, de transparencia, acceso a la información pública y buen gobierno⁶⁶⁷ (d'ara en endavant, Llei 19/2013). Aquests límits són d) la seguretat pública, g) les funcions administratives de vigilància, inspecció i control, i) la política econòmica i monetària, j) el secret professional i la propietat intel·lectual i industrial i k) la garantia de confidencialitat o el secret requerit en els processos de presa de decisió. Civio ha plantejat un recurs

⁶⁶³CONSEJO DE TRANSPARENCIA Y BUEN GOBIERNO, Resolución 701/2018, de 18 de febrero de 2019. Disponible a: <https://civio.app.box.com/s/0r9a58a10pdp2t08lnemlmc22sd7t31> Darrera visita el 2 de setembre de 2023.

⁶⁶⁴OBSERVATORI D'ÈTICA EN INTEL·LIGÈNCIA ARTIFICIAL DE CATALUNYA (OEIAC), *Intel·ligència artificial, ètica i societat...*, op. cit., 2021, pàg. 31.

⁶⁶⁵CONSEJO DE TRANSPARENCIA Y BUEN GOBIERNO, Contestación a la demanda en el Procedimiento Ordinario 18/2019 seguido en el Juzgado Central Contencioso-Administrativo núm. 8, de 5 de septiembre de 2019. Disponible a: <https://civio.app.box.com/s/q7b42raq1xxsrtjhg8owddnbjm8yn6ga> Darrera visita el 2 de setembre de 2023.

⁶⁶⁶ Juzgado Contencioso-Administrativo núm. 8, sentencia núm. 143/2021, de 30 de diciembre de 2021. Disponible a: <https://civio.app.box.com/s/ok4x3w0112iougk4cz6mt5ec73vdafb9> Darrera visita el 2 de setembre de 2023.

⁶⁶⁷ Ley 19/2013, de 9 de diciembre, de transparencia, acceso a la información pública y buen gobierno. BOE, núm. 295, 10 de diciembre de 2013. Espanya.

d'apel·lació i, per tant, el procediment segueix obert. En el sector privat ni tan sols es planteja la discussió.

Paral·lelament es posa en dubte que l'accés al codi font sigui una acció que aporti transparència, atès que sense coneixement tècnic no necessàriament aporta informació pertinent⁶⁶⁸. S'argumenta que l'accés no permet conèixer realment l'origen de la decisió si el biaix no està en el seu disseny sinó en el seu entrenament a través de dades⁶⁶⁹. Aquí és rellevant l'arquitectura del sistema algorítmic, que en el cas de ser més simple (p. e. determinar si una persona reuneix o no uns criteris preestablerts) l'accés al codi font o seqüència de programació aclarirà l'origen de la decisió, però en el cas d'algoritmes predictius més complexos on les dades d'entrenament adquireixen gran importància (p. e. algoritmes predictius emprats en la selecció de personal), no proporcionarà la base necessària per desafiar la discriminació patida arran del sistema algorítmic a cap persona que reclami⁶⁷⁰.

La transparència inclús pot ser insuficient quan, tot i possibilitar la identificació les persones responsables de les decisions automatitzades discriminatòries a qui exigir-los reparació del dany, qualsevol forma d'explicació i justificació de la seva conducta requereix la revelació de més informació sobre el sistema de presa de decisions⁶⁷¹. També pot tenir efectes col·laterals, ja sigui generant opacitat en submergir la informació rellevant entre molta de supèrflua, ja sigui instrumentalitzant-la en contra de grups i persones vulnerables⁶⁷².

Adicionalment, l'accés a informació en les disposicions que referenciem, és individual (ja que parteix del dret a la protecció de dades personals), la qual cosa sembla reduir l'efectivitat de la informació per fer front a les discriminacions algorítmiques.

⁶⁶⁸CASTETS-RENARD, C., "Accountability of Algorithms in the GDPR...", *op. cit.*, 2019, pàg. 106.

⁶⁶⁹RIVAS VALLEJO, P., "Discriminación algorítmica: detección, prevención y tutela", *op. cit.*, 2021, pàg. 9. L'autora defensa, així i tot, en un altre apartat del mateix article, que un hipotètic dret a accedir al codi font de l'algoritme en el context judicial podia salvar les objeccions plantejades en base al secret industrial a través de l'ajuda pericial, que podria garantir-lo al mateix temps que traduiria l'algoritme en el context del plet per aclarir la possible existència d'un biaix discriminatori (pàg. 47).

⁶⁷⁰*Idem*, pàg. 42.

⁶⁷¹ALMADA, M., "Human intervention in automated decision-making. Toward the construction of contestable Systems", *ICAIL '19: Proceedings of the Seventeenth International Conference on Artificial Intelligence and Law*, 2019, pàg. 10.

⁶⁷²ANANNY, M., i CRAWFORD, K., "Seeing without Knowing: Limitations of the Transparency Ideal and Its Application to Algorithmic Accountability", *New Media and Society*, vol. 20, núm. 3, 2018, pàg. 978 i 979.

Recapitulant, els articles 12.1, 13.2 (f), 14.2 (g) i 15.1 (h) i el considerant 71 de l'RGPD proporcionen la possibilitat a les persones que es veuen subjectes a una decisió automatitzada d'obtenir informació sobre ella. Malgrat això es presenten certs obstacles perquè aquesta informació sigui efectiva quan ens referim a decisions preses mitjançant sistemes algorítmics, com per exemple, que la informació que es dona només és individual, que interfereix amb els drets de propietat intel·lectual, que la complexitat inherent al sistema no es pugui explicar, que s'aconsegueixi accés a certa informació però que aquesta sigui inintel·ligible, o que es doni accés a certa informació, però que no serveixi per proporcionar proves en un cas de discriminació algorítmica.

Seguidament estudiem les mancances que trobem en els mecanismes de gestió del risc tal com establerts en l'RGPD per abordar la discriminació causada arran de sistemes algorítmics.

4. Límits a la gestió del risc de discriminació algorítmica per part de l'RGPD

El compliment d'allò que estableix l'RGPD exigeix cert nivell de maduresa dels sistemes de presa de decisions automatitzades emprats que permeti determinar de forma objectiva l'adequació dels tractaments i l'existència d'aval i mesures per gestionar els seus riscos. Per gestionar aquests riscos, les tecnologies de la intel·ligència artificial poden tenir un efecte multiplicador de les carències ètiques que ja estan presents en la societat o aquelles que s'arrossegueu del passat i estan registrades en dades històriques⁶⁷³.

En el cas SyRI, el Tribunal -malgrat declina avaluar si es compleixen les exigències de disseny i gestió del risc establertes per l'RGPD donada la manca d'informació-, subratlla que el sistema no preveu una revisió exhaustiva prèvia ni una revisió per part d'un tercer independent⁶⁷⁴, com tampoc una avaluació d'impacte relativa a la protecció de dades

⁶⁷³AEPD, *Adecuación al RGPD de tratamientos que incorporan Inteligencia Artificial. Una introducción*, Madrid, 2020, pàg. 49.

⁶⁷⁴Tribunal del Districte de La Haia, sentència de 5 de febrer de 2020, C/09/550982/ HA ZA 18-388, Case of SyRI, FJ 6.99.

(DPIA)⁶⁷⁵ per a cada projecte individual SyRI, que considera obligada donat l'abast i la gravetat de la invasió en la vida privada de les persones⁶⁷⁶.

La gestió del risc és el conjunt d'activitats i tasques que permeten controlar la incertesa relativa a una amenaça mitjançant una seqüència d'activitats que inclouen la identificació i l'avaluació del risc, així com mesures per a la seva reducció o mitigació⁶⁷⁷.

En aquest apartat estudiem les limitacions que trobem en la gestió del risc estipulada en l'àmbit de la protecció de dades per la protecció de les persones envers la discriminació algorítmica. En l'article 30 de l'RGPD⁶⁷⁸ es requereix un registre de les activitats de tractament i en l'article 32 de l'RGPD⁶⁷⁹ s'estableix la necessitat de definir i establir mesures de control i seguretat d'acord amb les particularitats de les activitats de tractament. Per una banda, orientat a la previsió del risc, l'RGPD planteja les avaluacions d'impacte de protecció de dades⁶⁸⁰, mecanisme basat en les avaluacions d'impacte a la privacitat⁶⁸¹, clàusules creades en règims de protecció de dades primerencs en els quals les organitzacions estaven obligades a registrar, notificar i contrastar amb les autoritats nacionals per assegurar l'adequació prèvia al tractament⁶⁸².

Quan ens referim a la presa de decisions mitjançant algorismes, un gran obstacle a l'efectivitat de l'avaluació d'impacte de protecció de dades és que es realitza de manera prèvia al tractament i no segueix l'evolució que pot tenir l'algoritme i el canvi d'ús que es pot fer fàcilment amb les dades (4.1). També, per avaluar el funcionament del sistema algorítmic, l'RGPD planteja en l'article 32.1. d) que la persona responsable i encarregada del tractament haurà de garantir sistemes de verificació, avaluació i valoració regulars en relació a la seguretat del tractament, és a dir, la realització d'auditories periòdiques. Si bé

⁶⁷⁵ La sigla prové de les paraules en anglès: Data Protection Impact Assessment.

⁶⁷⁶Tribunal del Districte de La Haia, sentència de 5 de febrer de 2020, C/09/550982/ HA ZA 18-388, Case of SyRI, FJ 6.105.

⁶⁷⁷AEPD, *Guía práctica de análisis de riesgos en los tratamientos de datos personales sujetos al RGPD*, Madrid, 2018, pàg. 3.

⁶⁷⁸Article 30 de l'RGPD.

⁶⁷⁹Article 32 de l'RGPD. Per una anàlisi d'aquest article en la doctrina espanyola, vegeu: ACINAS GARCÍA, N., "Seguridad de los tratamientos: los análisis de riesgos, la seudonimización, el cifrado y la adhesión a códigos de conducta (Comentario al artículo 32 RGPD)", a TRONCOSO REIGADA, A., (dir.) i VILLALBA CANO, L., (coord.), *Comentario al Reglamento General de Protección de Datos y a la Ley Orgánica de Protección de Datos Personales y Garantía de los Derechos Digitales*, Tomo I, Pamplona, Aranzadi, 2021, pàg. 2033-2042.

⁶⁸⁰En anglès: Data Protection Impact Assessments (DIPA's).

⁶⁸¹En anglès Privacy Impact Assessments (PIA's).

⁶⁸²BINNS, R., "Data Protection Impact Assessments: A Meta-Regulatory Approach", *International Data Privacy Law*, vol. 7, núm. 1, 2017, pàg. 23.

en la protecció de dades poden resultar efectives, l'opacitat i la intel·ligibilitat que comporten els sistemes algorítmics dificulten la seva funció (4.2).

4.1. L'avaluació d'impacte relativa a la protecció de dades

L'avaluació d'impacte⁶⁸³ és un mecanisme per identificar els riscos potencials, la probabilitat que aquests riscos apareguin i les conseqüències que poden tenir⁶⁸⁴. No està definida formalment en l'RGPD però l'article 35.7 especifica el seu contingut mínim⁶⁸⁵. Precisament una avaluació d'impacte de protecció de dades⁶⁸⁶, tal com està entesa en l'RGPD, serveix per subratllar els riscos per a la protecció de dades dels sistemes automatitzats que s'utilitzen per avaluar les persones basant-se en les seves dades personals. S'aplica en organitzacions públiques i privades, no es comparteix amb el públic i no té una revisió externa incorporada ni altres mecanismes individuals de garanties processuals⁶⁸⁷. Només és obligatòria sota certes circumstàncies on el tractament de dades és propens a resultar en un risc alt pels drets i llibertats de les persones⁶⁸⁸.

S'ha argumentat que les avaluacions d'impacte de la protecció de dades per a l'elaboració de perfils de dades de risc alt, conjuntament amb altres mesures preventives dins la protecció de dades, pot ser una bona manera d'abordar la discriminació algorítmica, ja que es restringeixen les accions de qui desplega els algoritmes⁶⁸⁹.

El considerant 71 de l'RGPD⁶⁹⁰ disposa que les companyies han de prevenir els efectes discriminatoris (entre altres tasques), però no estableix una manera per fer-ho. En canvi, les guies suggereixen a les companyies revisar si hi ha biaixos en les dades, revisar la

⁶⁸³ Per una anàlisi de les avaluacions d'impacte de protecció de dades em la doctrina espanyola, vegeu: MIRALLES LÓPEZ, R., "La evaluación de impacto relativa a la protección de datos (Comentario al artículo 35 RGPD)", a TRONCOSO REIGADA, A., (dir.) i VILLALBA CANO, L., (coord.), *Comentario al Reglamento General de Protección de Datos y a la Ley Orgánica de Protección de Datos Personales y Garantía de los Derechos Digitales*, Tomo I, Pamplona, Aranzadi, 2021, pàg. 2137-2162.

⁶⁸⁴ GREEN, B, "The False Promise...", *op. cit.*, 2020, pàg. 594.

⁶⁸⁵ Article 35.7 de l'RGPD.

⁶⁸⁶ En anglès: Data Protection Impact Assessments (DPIA's).

⁶⁸⁷ REISMAN, D., SCHULTZ, J., CRAWFORD, K. i WHITTAKER, M., *Algorithmic Impact Assessments: a Practical Framework...*, *op. cit.*, 2018, pàg. 7.

⁶⁸⁸ Article 35 de l'RGPD.

⁶⁸⁹ VEALE, M. i EDWARDS, L., "Slave to the Algorithm? Why a 'Right to an Explanation' Is Probably Not the Remedy You Are Looking For", *Duke Law and Technology Review*, vol. 16, núm. 1, 2017, pàg. 77-80.

⁶⁹⁰ Considerant 71 de l'RGPD.

precisió i rellevància de les decisions, desenvolupar sistemes d'auditories dels algorismes i usar procediments i mesures per prevenir errors, imprecisions i discriminació en base a categories protegides⁶⁹¹. L'RGPD identifica el problema, suggereix què han de considerar les persones reguladores adequat, i sovint requereix a les companyies que cooperativament busquin solucions. Les avaluacions d'impacte de protecció de dades - atès que podrien reunir totes les salvaguardes per tecnologies específiques-, podrien jugar un rol potencial en l'establiment de normes⁶⁹².

L'RGPD no inclou explícitament perspectives precises sobre com i a quin nivell de detall l'impacte en els drets fonamentals ha de ser avaluat. En el context de les decisions automatitzades, la persona ha de ser informada sobre la rellevància i les conseqüències previstes d'aquesta decisió⁶⁹³. En el considerant 75, s'estableix que riscos potencials als drets, com la discriminació o qualsevol altre desavantatge econòmic o social també poden formar part de l'avaluació⁶⁹⁴.

Així i tot, també s'han fet paleses les limitacions d'aquest mecanisme. En contextos en els quals les avaluacions d'impacte de protecció de dades han esdevingut obligatòries, s'han descrit com a sistemes massa prescriptius i formulistes, buits de qualsevol contingut rellevant per a la protecció de dades, més enllà dels requeriments legals⁶⁹⁵ (de la mateixa manera, Martínez Martínez considera que el compliment en la protecció de dades segueix essent bàsicament formal⁶⁹⁶). A més, Jasanoff els considera una exemplificació de la percepció generalitzada dels reptes socials com a tecnològics per naturalesa i adaptables a solucions tecnològiques⁶⁹⁷, fent parar l'atenció sobre el seu caràcter profundament normatiu, mostrant que al contrari de ser un mètode objectiu, es tracta d'un mitjà molt particular d'enquadrar percepcions, reduir anàlisi, esborrar incerteses i desactivar les

⁶⁹¹ ARTICLE 29 DATA PROTECTION WORKING PARTY, *Guidelines on Automated individual...*, *op. cit.*, 2018, pàg. 28.

⁶⁹² ROIG, A., "Safeguards for the Right Not To Be Subject to a Decision Based Solely on Automated Processing (Article 22 GDPR)", *European Journal of Law and Technology*, vol. 8, núm. 3, 2017, pàg. 2.

⁶⁹³ Article 13.2 (f), 14.2 (g) i 15.1 (h) de l'RGPD.

⁶⁹⁴ Considerant 75 de l'RGPD.

⁶⁹⁵ BINNS, R., "Data Protection Impact Assessments...", *op. cit.*, 2017, pàg. 26.

⁶⁹⁶ MARTÍNEZ MARTÍNEZ, R., "Una aproximación crítica al Reglamento General de Protección de Datos", *Diario La Ley*, núm. 10294, Sección Tribuna, 25 de Mayo de 2023. Disponible a: https://diariolaley.laleynext.es/Content/Documento.aspx?params=H4sIAAAAAAAAAEAE1QTWvDMAz9NfMIMLJIFHrwJQ2UQRmly3ZXbS0xc6zMlrPk309JGZtBWNL74NlfGePS4sz6LUABY6TZDWDcXS7LjyoUJq4dWnZGYF9csPMwYGAqjhgwyspicY7EaP5UsmqAKam0BArLoNuYUTFck66eFBjO4BsyeuvdhC1cV4CixVgvulRMDP6CST887IXq6fsFJtcBOwoIxJudsIY3bSmm2u135U5NGJMQ9LvrJCCq3nX9SYpv_IQQTX-GDvXJQaxz57BUrqHNM7Kh0_J9Lqx_ivqzCyuVw43TBkvt7wPD-Ax2N88MI5-uZCXINs8UhK_vH7WczhApJzQ6_IH3dTnGnIBAAA=WKE Darrera visita el 2 de setembre de 2023.

⁶⁹⁷ GREEN, B., "The False Promise...", *op. cit.*, 2020, pàg. 600.

polítiques⁶⁹⁸. No només es tracta de posar en qüestió l'activitat de processament de dades, sinó que cal tenir presents els altres efectes que es posaran de relleu posteriorment en el comportament individual de les persones, i la interacció política i social⁶⁹⁹. Per això, un repte de les avaluacions d'impacte que es puguin fer en els sistemes algorítmics és l'apreciació dels possibles danys culturals i socials. El problema és la necessitat de fer assumpcions o prediccions sobre factors socials i culturals que varien enormement entre comunitats i àrees geogràfiques, i aquesta pràctica sovint resulta en troballes que només reflecteixen possibles impactes en la cultura dominant i ometen o interpreten malament els impactes en comunitats o persones marginades⁷⁰⁰.

En relació als biaixos que també poden contenir les avaluacions d'impacte i les discriminacions que els poden passar per alt, un exemple que il·lustra de la seva ineficàcia en relació als col·lectius vulnerables i oprimits de la societat és el de les avaluacions d'impacte que s'apliquen a persones que presumptivament han comès un delictes amb la finalitat de determinar la seva perillositat i si han d'ingressar a presó preventiva. Tot i que el seu objectiu sigui eliminar els biaixos subjectius dels/de les jutges/sses, les avaluacions d'impacte també estan incloses en el sistema, en el qual incentius estructurals i polítics afavoreixen en gran mesura les polítiques punitives, proclius a legitimar la lògica del càstig penal, i, de vegades, inefectives perquè es permet als/les jutges/sses fer ús de la seva discrecionalitat per ignorar-les o usar-les de manera selectiva, segons la informació o recomanacions que proveeixin⁷⁰¹. Green critica que malgrat el suport generalitzat que tenen, es basen en una teoria deficient del canvi, sense proveir objectivitat ni capacitat de reforma, ja que sovint hi ha molts judicis subjectius, assumint que les característiques de la població són constants, tal com que certs factors que van produir certs resultats en el passat, produiran els mateixos resultats en el futur; legitimant i naturalitzant lògiques de presó (p. e. als investigats perillosos se'ls ha d'ingressar en presó preventiva) i pràctiques (p. e. com determinar que els investigats són perillosos); i basades en concepcions

⁶⁹⁸JASANOFF, S., *Designs on Nature: Science and Democracy in Europe and the United States*, Princeton, New Jersey, Princeton University Press, 2011, pàg. 266.

⁶⁹⁹RAAB, C. D., "Information privacy, impact assessment, and the place of ethics", *Computer Law and Security Review*, vol. 37, 2020, pàg. 11.

⁷⁰⁰REISMAN, D., SCHULTZ, J., CRAWFORD, K. i WHITTAKER, M., *Algorithmic Impact Assessments: a Practical Framework...*, *op. cit.*, 2018, pàg. 18.

⁷⁰¹GREEN, B., "The False Promise...", *op. cit.*, 2020, pàg. 597-599.

individualistes de risc i biaix que porten a remeis individualistes i no eficaces en relació a la discriminació racial i a l'empresonament massiu⁷⁰².

S'evidencia així, com també pot passar amb les avaluacions d'impacte aplicables a la protecció de dades, que poden generar-se biaixos i problemàtiques estructurals de la mateixa manera que el tractament en si, i que inclús poden participar a donar continuïtat a conductes a la pràctica discriminatòries. De manera que avaluacions d'impacte considerades "justes", el que fan es subratllar la seva ceguesa en una societat desigual⁷⁰³. Diagnosticant els biaixos com a comportaments dels individus típicament no intencionals, ignoren les polítiques i institucions que estructurin les jerarquies socials. La discriminació i l'opressió estan produïdes no només per persones que fan judicis esbiaixats, sinó també a través de lleis i institucions que sistemàticament beneficien un grup sobre un altre⁷⁰⁴. Com a eines per detectar i mitigar la discriminació, poden esdevenir caselles de verificació que les entitats marquen i obliden, deixant de banda potencialment les preocupacions comunitàries⁷⁰⁵.

Mantelero també planteja que per fer front als impactes potencials de la IA, que afecten a diversos drets i llibertats, s'han de repensar les metodologies d'avaluació, que haurien d'anar més enllà de la perspectiva limitada de les avaluacions d'impacte de protecció de dades, que no consideren adequadament els drets potencialment afectats, la seva diversitat i complexitat o els conflictes ètics i socials⁷⁰⁶.

Finalment, i des d'una vessant processal, el que s'ha criticat a les avaluacions d'impacte de protecció de dades amb més insistència, i que té més rellevància en relació amb la discriminació, és el fet que no es requereixi legalment que es revelin al públic, ni quan s'han finalitzat (està altament recomanat, si més no en forma de resums significatius, però

⁷⁰²Ídem, pàg. 597-602.

⁷⁰³Ídem, pàg. 601.

⁷⁰⁴Ídem, pàg. 599.

⁷⁰⁵REISMAN, D., SCHULTZ, J., CRAWFORD, K. i WHITTAKER, M., *Algorithmic Impact Assessments: a Practical Framework...*, op. cit., 2018, pàg. 7-10. Vegeu, també: OBASOGIE, O. K., "The return of Biological Race? Regulating Race and Genetics through Administrative Agency Race Impact Assessments", *Southern California Interdisciplinary Law Journal*, vol. 22, núm. 1, 2012, pàg. 59. L'autor especifica que sense suport normatiu, les avaluacions d'impacte poden esdevenir una eina únicament processal que no sigui capaç de crear cap canvi.

⁷⁰⁶ MANTELERO, A., *Beyond Data. Human Rights, Ethical and Social Impact Assessment in AI*, The Hague, Asser Press, 2022, pàg. 12.

no és obligat)⁷⁰⁷, generant que el mecanisme perdi eficiència, i impeding la possibilitat d'incloure una revisió experta externa o un procés de notificació i audiència per la revisió pública⁷⁰⁸. Una avaluació d'impacte usualment implica tres parts: l'entitat regulada, qui regula i terceres parts que es veuen impactades o organitzacions de la societat civil. En l'RGPD s'empra internament o com a molt en conversa amb qui regula⁷⁰⁹, de manera que un element essencial en l'eina⁷¹⁰, que permet un diàleg amb la població general⁷¹¹, està completament absent.

L'altre sistema principal de gestió del risc que planteja l'RGPD és la realització d'auditories. La seva eficiència en relació a la protecció de la discriminació en les decisions automatitzades és el tema que tractarem a continuació.

4.2. Sistemes d'auditories insuficients

La implementació de sistemes que afecten a la població sense marcs per la rendició de comptes suposa el risc de perdre control sobre com s'han pres les decisions, dificultant la identificació i actuació contra els biaixos, errors o altres problemes⁷¹². No obstant això, l'existència de sistemes que realment no siguin eficaços i presentin una situació de control il·lusòria també esdevenen un problema, ja que asseguren un control que realment no existeix (per exemple, tal com estan establertes en l'RGPD, les auditories són confiades a les pròpies entitats i organitzacions que fan el tractament).

⁷⁰⁷ ARTICLE 29 DATA PROTECTION WORKING PARTY, *Guidelines on Data Protection Impact Assessment (DPIA) and determining whether processing is "likely to result in a high risk" for the purposes of Regulation 2016/679*, 17/EN, WP 248 rev.01, Brussels, 4 October 2017, pàg. 18.

⁷⁰⁸ REISMAN, D., SCHULTZ, J., CRAWFORD, K. i WHITTAKER, M., *Algorithmic Impact Assessments: a Practical Framework...*, *op. cit.*, 2018, pàg. 7.

⁷⁰⁹ KAMINSKI, M.E., "Binary Governance...", *op. cit.*, 2019, pàg. 1609. Per altres autors/es que reclamen la supervisió per part de la societat civil, vegeu: VILLANI, C., *Donner un sens à l'intelligence artificielle...*, *op. cit.*, 2018, pàg. 144; IVANOVA, Y., "The Data Protection Impact Assessment as a Tool to Enforce Non-discriminatory AI", *Privacy Technologies and Policy: 8th Annual Privacy Forum*, 2020, pàg. 11.

⁷¹⁰ Vegeu, per exemple: REISMAN, D., SCHULTZ, J., CRAWFORD, K. i WHITTAKER, M., *Algorithmic Impact Assessments: a Practical Framework...*, *op. cit.*, 2018, pàg. 9 i s.

⁷¹¹ KAMINSKI, M. E. i MALGIERI, G., "Algorithmic impact assessments under the GDPR: producing multi-layered explanations", *International Data Privacy Law*, vol. 11, núm. 2, 2021, pàg. 133.

⁷¹² REISMAN, D., SCHULTZ, J., CRAWFORD, K. i WHITTAKER, M., *Algorithmic Impact Assessments: a Practical Framework...*, *op. cit.*, 2018, pàg. 4.

L'article 32.1 de l'RGPD disposa que la persona responsable i encarregada del tractament aplicarà mesures tècniques i organitzatives adequades per garantir el nivell de seguretat apropiat al risc, incloent, entre d'altres, la pseudonimització i el xifratge de les dades personals; la capacitat de garantir la confidencialitat, integritat, disponibilitat i resiliència permanents; la capacitat de restaurar la disponibilitat i l'accés a les dades personals de forma ràpida en cas d'incident físic o tècnic; i finalment en el seu apartat d), un procés de verificació, avaluació i valoració regulars de l'eficàcia de les mesures que vetllen per assegurar el tractament⁷¹³. Aquest últim procés té com a objectiu disminuir els riscos que comporta el tractament de dades, millorar-ne el funcionament⁷¹⁴, i documentar i recollir informació suficient per permetre de forma satisfactòria i demostrable, acreditar les accions preses⁷¹⁵. Nogensmenys, els sistemes de presa de decisions automatitzades administratius es critiquen perquè sovint fallen a retenir qualsevol registre d'auditoria de com van prendre la decisió o sobre quines dades la decisió es va basar⁷¹⁶. Si no se sap en base a quines dades es va prendre la decisió, localitzar com i per què s'ha pres una decisió discriminatòria esdevé quasi impossible.

Les decisions automatitzades s'utilitzen en un context determinat, amb un tractament de dades específic, amb unes característiques singulars i en un entorn d'explotació determinat. Les auditories⁷¹⁷ (que poden ser internes o externes) han de tenir en compte el context i l'entorn, ja que de no ser així oferiran resultats parcials que no s'adequaran a la realitat. En aquest sentit, la homogeneïtat del tractament de dades, desconeixent les circumstàncies personals i particulars de cada persona, és de difícil adaptació amb el principi d'igualtat. Un dels aspectes claus que ha de prioritzar una auditoria és garantir que el sistema s'està usant pel propòsit pel qual va ser dissenyat. Sobretot, quan es tracta d'un component adquirit per un tercer, s'ha de determinar quines altres funcions

⁷¹³Article 32.1 de l'RGPD.

⁷¹⁴Es fa referència a la realització d'auditories en l'article 28.1 (h), 39.1 (b), 47.2 (g) i 58.1 (b).

⁷¹⁵AEPD, *Adecuación al RGPD de tratamientos que incorporan Inteligencia Artificial. Una introducción*, Madrid, 2020, pàg. 45.

⁷¹⁶CRAWFORD, K. i SCHULTZ, J., "Big data and due process...", *op. cit.*, 2014, pàg. 124.

⁷¹⁷ Per una anàlisi de les auditories a l'RGPD a la doctrina espanyola, vegeu: DEL PESO NAVARRO, E. i DEL PESO RUIZ, M., "El control interno y la auditoría de los tratamientos en el marco del RGPD (Comentario al artículo 32 RGPD)", a TRONCOSO REIGADA, A., (dir.) i VILLALBA CANO, L., (coord.), *Comentario al Reglamento General de Protección de Datos y a la Ley Orgánica de Protección de Datos Personales y Garantía de los Derechos Digitales*, Tomo I, Pamplona, Aranzadi, 2021, pàg. 2043-2067.

col·laterals està realitzant aquest component i si hi poden haver repercussions normatives.⁷¹⁸

S'erigeixen dubtes en relació a l'efectivitat de les auditories o les intervencions de persones per retre comptes a causa de la manca de traçabilitat i visibilitat, i en quin grau brinden protecció enfront les discriminacions. Les persones que intervenen no tenen autonomia completa, ja que només supervisen allò que indiquen les categoritzacions de l'algoritme⁷¹⁹. Com els sistemes estan basats en elements inaccessibles a les persones (per exemple, anàlisi de grans bases de dades), es requereix per part de la persona cert nivell de confiança en els càlculs dels algoritmes⁷²⁰. Això comporta també que les persones afectades perdin la seva capacitat de desafiar les decisions preses pel sistema⁷²¹.

Així i tot, contràriament a les persones, els raonaments de les quals poden ser ocults als altres (és a dir, també són caixes negres, ja que el seu procés de decisió no es pot gravar, replicar o auditar), els algoritmes, el funcionament dels quals pot ser de la mateixa manera misteriós (no podem saber com les dades d'entrada es converteixen en resultats), també poden ser dispositius de gravació com els sistemes de monitorització de dades dels avions, trens i cotxes, sorgint l'oportunitat -que no sorgeix amb les persones-, de gravar, replicar i auditar el procés que segueixen⁷²². Kroll *et al.* defensen que l'auditoria és una estratègia tècnicament limitada per abordar els biaixos en els sistemes algorítmics i que s'ha d'apostar per implementar eines tècniques⁷²³, advocant per implementar una perspectiva *ex ante*. Kim respon que malgrat les limitacions, les auditories per detectar la discriminació haurien de restar com una part important de l'estratègia per detectar i abordar els biaixos algorítmics, ja que les eines tècniques únicament no poden, de manera fiable, prevenir resultats discriminatoris que habitualment són causats per un procés social més ampli i no únicament del codi font de l'algoritme, per tant, implementar les millors eines tècniques disponibles mai pot garantir que no hi ha biaixos en els algoritmes⁷²⁴.

⁷¹⁸AEPD, *Adecuación al RGPD de tratamientos que incorporan Inteligencia Artificial. Una introducción*, Madrid, 2020, pàg. 46.

⁷¹⁹ LEESE, M., "The New Profiling: Algorithms, Black Boxes...", *op. cit.*, 2014, pàg. 505.

⁷²⁰ *Ídem, ibidem*.

⁷²¹ *Ídem, ibidem*.

⁷²² PASQUALE, F., *The Black box society: The secret algorithms that control Money and information*, *op. cit.*, 2015, pàg. 3. És en aquest treball en el qual l'autor planteja que la metàfora de la caixa negra s'aplica en ambdós sentits. No se sap què hi ha a l'interior però tot queda registrat.

⁷²³ KROLL, J.A. *et al.*, "Accountable algorithms...", *op. cit.*, 2017, pàg. 694.

⁷²⁴ KIM, P. T., "Auditing Algorithms for Discrimination", *University of Pennsylvania Law Review Online*, vol. 166, 2017, pàg. 190 i 191.

Aquesta discussió sembla un debat entre l'eficiència d'una perspectiva reguladora *ex post* (auditoria) davant la possibilitat d'una perspectiva reguladora *ex ante* (eines tècniques). Per abordar la discriminació algorítmica sembla necessari implementar les dues perspectives i de manera cíclica. No només s'ha de treballar per implementar eines tècniques des del disseny sinó que també cal realitzar auditories per valorar l'impacte dels algoritmes. Les auditories, per si soles, no semblen ser efectives i, com planteja Kim⁷²⁵, han de restar com un element essencial però han de formar part d'un conjunt de mecanismes.

En definitiva, les limitacions principals que trobem en les eines estudiades -avaluacions d'impacte de protecció de dades i auditories- per ser útils en referència a la discriminació algorítmica, són la dificultat que puguin aplicar-se a sistemes que evolucionen amb el temps (com poden ser els sistemes algorítmics de presa de decisió causants de discriminació). És a dir, si un tractament de dades té un inici i un final, en el cas d'emprar un sistema algorítmic capaç d'evolucionar amb el temps, el sistema que ha iniciat el tractament pot no ser el mateix que l'hagi acabat; i pot ser que els paràmetres que defineixin aquestes eines tinguin els mateixos biaixos (explícits i implícits) que es poden trobar en el tractament en si, amb el risc d'esdevenir processos merament formalistes i normatius sense més rellevància.

Tot seguit analitzem la perspectiva individual de la protecció que proposa l'RGPD i com pot afectar a la protecció de la discriminació algorítmica.

5. Perspectiva individual de la protecció que confereix l'RGPD

L'anàlisi massiva de dades fa possible recollir i analitzar grans quantitats de dades per intentar identificar patrons en el comportament de grups i individus, prenent decisions que afecten les dinàmiques internes del grup, amb conseqüències per a les problemàtiques col·lectives de les persones involucrades⁷²⁶. En aquest apartat estudiem, en primer lloc,

⁷²⁵ *Ídem*, pàg. 202.

⁷²⁶ MANTELERO, A., "Personal data for decisional purposes in the age of analytics: From an individual to a collective dimension of data protection", *Computer Law & Security Review*, vol. 32, núm. 2, 2016, pàg. 245.

les limitacions de la perspectiva marcadament individual de l'RGPD per contribuir en la protecció envers discriminació algorítmica (5.1), i, en segon lloc, més específicament, les limitacions dels drets individuals que atorga l'article 22.3 de l'RGPD (5.2).

5.1. Limitacions de la perspectiva individual de la protecció

La protecció de les persones en matèria de decisions automatitzades s'aborda des de la perspectiva individual, ja que els referents normatius que es tenen en compte per regular-les, a saber, el dret a la privacitat, o el més recent dret a la protecció de dades, estan protegits com a drets individuals⁷²⁷, imperant la concepció de la propietat com a individual. Podem detectar certes limitacions que presenta aquesta perspectiva quan ens referim a la discriminació causada per les decisions automatitzades.

En primer lloc, l'anàlisi massiva de dades sembla difícil d'enquadrar en els drets de propietat individual. Primerament, perquè la decisió resultant del tractament de dades personals que té origen en només una persona té un potencial impacte en altres⁷²⁸, de manera que el control individual sobre les dades personals i l'exercici dels drets de propietat individuals poden ser difícils d'aconseguir, i segonament, perquè crear i gestionar drets de propietat que siguin transparents esdevé també un repte. Això és a causa, segons Purtova (i com ja hem tractat tant en l'apartat 3.1.1 del capítol I i 2.2 d'aquest capítol), de la perspectiva adoptada a Europa de la definició de "dades personals": les mateixes dades, depenent del context, poden ser personals o no personals⁷²⁹. Per tant, generar i predir on i quan s'han d'establir proteccions de privacitat sobre les dades esdevé molt complicat, ja que no es pot saber per avançat quan un algoritme d'aprenentatge automàtic pot predir informació personal identificable sobre una persona⁷³⁰.

En segon lloc, la presa d'una decisió discriminatòria individual normalment té en consideració múltiples criteris perquè es categoritza a les persones (per exemple: color de

⁷²⁷ *Ídem*, pàg. 238.

⁷²⁸ PURTOVA, N., "Do property rights in personal data...", *op. cit.*, 2017, pàg. 72.

⁷²⁹ *Ídem*, pàg. 64.

⁷³⁰ CRAWFORD, K. i SCHULTZ, J., "Big data and due process...", *op. cit.*, 2014, pàg. 99.

pell, edat, gènere). S'adopta així una dimensió col·lectiva, ja que la manera de classificar dels algoritmes no se centra merament en els individus, sinó en grups de persones amb les mateixes característiques. La preocupació més imperant en aquest context és la protecció d'aquests grups del perjudici potencial causat per l'anàlisi de dades invasiu i discriminatori⁷³¹, en conseqüència, es tracta d'un problema global que concerneix no només a una persona sinó de manera més general, a grups de persones que representen poblacions vulnerables⁷³² (com ja hem vist en l'apartat 2.2 del capítol II). Es mostren llavors insuficients els drets individuals de protecció de dades, incrementant els arguments a favor d'una privacitat de grup o alguna forma de drets col·lectius⁷³³ que respongui a qüestions com: quina protecció es brinda a les decisions automatitzades que impliquen un grup? Què passa amb aquelles persones que es veuen col·lateralment afectades a causa d'una decisió individual automatitzada? Totes les decisions automatitzades, a causa de la categorització, impliquen un grup? Partint de la base que les discriminacions ocasionades per decisions individuals automatitzades no estan suficientment protegides i que les que afecten a grups de persones directament no estan contemplades, no seria pertinent plantejar-se si una protecció col·lectiva dels drets milloraria la protecció de totes les persones? Aquestes qüestions les reprenem en la tercera part d'aquest treball.

En tercer lloc, les persones no saben que pertanyen a un grup creat pels sistemes de recollida de dades, no són conscients de la identitat d'altres membres del grup, no tenen cap relació amb ells i tenen una percepció limitada de les seves problemàtiques col·lectives. A més a més, aquests grups creats per l'anàlisi de les dades tenen una geometria variable i els individus poden canviar d'un grup a un altre, al contrari dels grups tradicionalment entesos on hi ha una consciència de formar part d'un grup i on hi ha visibilitat externa⁷³⁴. En aquest sentit, des d'una perspectiva sociològica, els membres d'un grup són conscients que estan essent part d'un grup i el grup és normalment reconegut com una estructura social autònoma. En base a aquesta posició, un grup és un producte de decisions concurrents de varies persones que lluiten per trobar un objectiu comú o compartir experiències comunes, valors o interessos⁷³⁵. Aquesta dinàmica és

⁷³¹MANTELERO, A., "Personal data for decisional purposes...", *op. cit.*, 2016, pàg. 246.

⁷³²CASTETS-RENARD, C., "Accountability of Algorithms in the GDPR...", *op. cit.*, 2019, pàg. 125.

⁷³³MANTELERO, A., "Personal data for decisional purposes...", *op. cit.*, 2016, pàg. 254.

⁷³⁴*Idem*, pàg. 245-246. Per aprofundir en la inexistència, en l'espai digital, d'un grup capaç d'atacar les relacions existents de dominació, vegeu: HAN, B-C., *En el enjambre*, Barcelona, Herder, 2021, pàg. 27-32.

⁷³⁵*Idem*, pàg. 244.

impossible d'aplicar a les decisions automatitzades a causa de la manca de l'element principal perquè sigui possible el reconeixement de l'existència del grup.

A continuació abordem les limitacions que trobem en l'exercici dels drets que confereix a les persones l'article 22.3 de l'RGPD quan són subjectes d'una decisió únicament automatitzada i les dificultats que presenten per conferir un sistema d'oposició efectiu envers la discriminació algorítmica si no es complementen amb altres eines.

5.2. Drets individuals de l'article 22.3 de l'RGPD

L'article 22 de l'RGPD permet als subjectes de decisions basades únicament en tractament automatitzat, d'accedir a un règim d'oposició, consistent en l'exercici dels tres drets⁷³⁶, a saber, el dret a la intervenció humana (5.2.1), el dret a impugnar la decisió basada únicament en el tractament automatitzat (5.2.2) i el dret a expressar el seu punt de vista (5.2.3).

De manera general, es plantegen tres entrebancs per poder posar-los en pràctica. En primer lloc, a causa de factors com la manca d'informació sobre el sistema o la indecisió sobre si hi ha hagut o no participació humana significativa, la persona que cedeix les seves dades i s'ha vist discriminada té molt difícil l'exercici d'aquest drets. Per exemple, en el cas SyRI, el Tribunal raona que donada la impossibilitat de verificar el procés de decisió i els passos que segueix el sistema per produir un informe de risc, és difícil d'entendre, en primer lloc, com la persona que cedeix les seves dades podrà defensar-se davant el fet que es produeixi un informe de risc sobre ella, i en segon lloc, com en el cas que no s'hagi acabat generant un informe de risc sobre la persona, aquesta podrà saber si les seves dades s'han tractat correctament⁷³⁷. Per tal que els drets individuals de l'RGPD tinguin incidència en l'àmbit de la discriminació, la persona necessita no només informació sobre la decisió que l'afecta, sinó també informació sobre com l'algoritme tracta diferents grups

⁷³⁶AMRAM, D., "The Role of the GDPR in Designing the European Strategy on Artificial Intelligence: Law-Making Potentialities of a Recurrent Synecdoche", *Opinio Juris in Comparatione. Studies in comparative and National Law*, vol. 1, núm. 1, 2020, pàg. 84.

⁷³⁷Tribunal del Districte de La Haia, sentència de 5 de febrer de 2020, C/09/550982/ HA ZA 18-388, Case of SyRI, FJ 6.90.

(que no és el mateix que la informació d'altres persones), i la tendència de l'algoritme cap al biaix i la discriminació⁷³⁸.

També, la manca de mitjans per fer efectiu un canvi o una càrrega de treball excessiva pot ocasionar que, malgrat la persona que intervingui tingui el poder de canviar els resultats d'un sistema automatitzat, sigui incapaç d'exercir-lo a la pràctica, inclús essent responsable dels resultats adversos⁷³⁹. Pot esdevenir, llavors, un repte identificar si realment la participació humana ha estat significativa o no, per tant, dubtar si l'automatització ha estat parcial o total, ocasionant que les persones vegin bloquejada l'oportunitat d'exercir els seus drets.

En segon lloc, aquests drets són individuals, i només poden ser emprats quan es tracta de decisions que es basen, almenys en part, en dades que es refereixen a la persona que els exercita. Aquesta perspectiva individuocèntrica exclou de l'àmbit del dret algunes formes de presa de decisió automatitzada, com per exemple, decisions que es basen únicament en dades anonimitzades o dades agregades⁷⁴⁰ o decisions que es basen en inferències extrems de dades personals⁷⁴¹. També, quan s'elaboren perfils de grup o s'apliquen com a resultat de dades que no es poden relacionar amb individus, la protecció de dades no serveix (per exemple, després de recollir les dades, aquestes poden ser anonimitzades, de manera que la protecció de dades ja no s'aplica⁷⁴²).

En tercer lloc, la problemàtica més evident resideix en la seva efectivitat i la seva capacitat per crear un sistema d'oposició il·lusori. Com exposa Gamper posant d'exemple les queixes pel funcionament discriminatori en el si de les institucions, la manifestació de la queixa permet que la persona es desfogui i d'aquesta manera el propi sistema s'autolegitima, contenint i neutralitzant la protesta⁷⁴³. Addicionalment, l'existència dels drets només alleugereix marginalment la seva posició de desavantatge. Les dades personals d'una persona poden ser tractades per molts responsables de tractament de

⁷³⁸ KAMINSKI, M. E. i MALGIERI, G., "Algorithmic impact assessments...", *op. cit.*, 2021, pàg. 143.

⁷³⁹ WAGNER, B., "Liable, but Not in Control? Ensuring Meaningful Human Agency in Automated Decision-Making Systems", *Policy & Internet*, vol. 11, núm. 1, 2019, pàg. 111.

⁷⁴⁰ BRKAN, M., "Do Algorithms Rule the World? Algorithmic Decision-Making and Data Protection in the Framework of the GDPR and Beyond", *International Journal of Law and Information Technology*, vol. 27, núm. 2, 2019, pàg. 100.

⁷⁴¹ WACHTER, S. i MITTELSTADT, B., "A Right to Reasonable Inferences: Re-Thinking Data Protection Law in the Age of Big Data and AI", *Columbia Business Law Review*, núm. 2, 2019, pàg. 498.

⁷⁴² SCHREURS, W., "Cogitas, Ergo Sum. The role of data protection...", *op. cit.*, 2008, pàg. 245.

⁷⁴³ GAMPER, D., *Las mejores palabras. De la libre expresión*, Barcelona, Anagrama, 2019, pàg. 100.

moltes maneres diferents i l'RGPD no li concedeix eines per mantenir i controlar els responsables i les activitats de tractament⁷⁴⁴.

Específicament, cal abordar les dificultats que trobem en la protecció que pot brindar cada dret referit en l'article 22.3 de l'RGPD per lluitar envers la discriminació algorítmica.

5.2.1. Dret a la intervenció humana

L'article 22.3 de l'RGPD exigeix a la persona responsable del tractament d'implantar un procediment que en el seu cas permeti que la persona afectada pugui arribar a interactuar amb un membre pertanyent a l'organització que ha dut a terme el tractament⁷⁴⁵. D'aquest dret se'n desprenen els altres dos, ja que si no hi ha una persona no podem expressar el nostre punt de vista a ningú ni impugnar la decisió de manera efectiva.

En l'exercici del dret, tota supervisió de la decisió ha de ser duta a terme per una autoritat competent amb la capacitat de canviar-la -de no ser així, aquest dret seria inútil-, ha de ser significativa i ha de tenir en compte totes les dades rellevants, de manera que no sigui una mera formalitat. D'aquesta afirmació es deriven dues qüestions, a saber, d'una banda sembla inversemblant que quan ens referim al processament massiu de dades la intervenció humana pugui ser mai suficient i satisfactòria, i d'altra banda, fins a quin punt la intervenció humana pot ser més perjudicial que beneficiosa quan pot ser la que aporti una visió esbiaixada.

En referència al primer supòsit, podem exemplificar la insuficiència d'aquest dret en el fet que donar la possibilitat a la intervenció humana es limita a l'oportunitat d'interactuar amb un ésser humà i no amb una màquina, però no s'assegura que es proveeixi d'una millor comprensió⁷⁴⁶, o que, si bé es podrà verificar que el funcionament del procés, l'algoritme o la màquina sigui el correcte, no es podrà verificar que les dades emprades de cada individu per arribar a unes conclusions determinades no siguin esbiaixades. Ho

⁷⁴⁴WOLTERS, P.T.J., "The Control by and Rights of the Data Subject Under the GDPR", *Journal of Internet Law*, vol. 22, núm. 1, 2018, pàg. 16.

⁷⁴⁵PALMA ORTIGOSA, A., "Decisiones Automatizadas en el RGPD. El uso de algoritmos en el contexto de la protección de datos" *Revista General de Derecho Administrativo*, núm. 50, 2019, pàg. 33.

⁷⁴⁶CASTETS-RENARD, C., "Accountability of Algorithms in the GDPR...", *op. cit.*, 2019, pàg. 121.

podrà fer l'algoritme, però no ho podrà fer una persona, a causa de la gran quantitat d'informació que s'estarà tractant. En aquest sentit, en el cas SyRI, el tribunal considera insuficient que el procediment d'avaluació de la validesa de les dades desxifrades per la investigació inclogui la verificació per part d'una persona dels falsos positius i falsos negatius, ja que la forma en què es fa la selecció definitiva del risc no és pública⁷⁴⁷, de manera que la intervenció humana no és suficientment significativa.

En referència al segon supòsit, podem exemplificar-lo en la intervenció no beneficiosa que poden tenir els éssers humans com a supervisors, atès que pot passar que les persones que intervenen siguin més proclius a actuar de maneres discriminatòries contra individus o grups⁷⁴⁸ que no els sistemes en si.

Si bé és cert que la intervenció humana *post hoc* es pot entendre com una protecció en contra d'algoritmes intencionadament perjudicials i com una evidència que fins i tot el software dissenyat en base a les millors pràctiques tècniques pot fallar en casos específics⁷⁴⁹, també hem de tenir en compte que la persona amb la qual interactuarem no serà contractada per una autoritat independent sinó per l'organització pública o privada que tingui interès en què es prengui la decisió i, per tant, existirà sempre una asimetria de poder (la persona supervisora sempre tindrà més informació, més mitjans i més recursos). En aquest sentit, també en el cas SyRI, el Tribunal argumenta que el Govern dels Països Baixos no proporciona cap explicació sobre la protecció de la divisió funcional interna de les unitats d'inspecció dels informes de risc, ja que l'òrgan consultiu encarregat de validar-los està integrat per representants d'òrgans que també tenen interès en combatre i prevenir abusos i fraus⁷⁵⁰.

Finalment, l'evidència que els sistemes algorítmics puguin analitzar quantitats de dades inabastables pels éssers humans, és a dir, puguin fer alguna cosa que les persones no poden fer, deixa a aquestes últimes en una posició de desavantatge per prendre una decisió no automatitzada diferent a la decisió automatitzada suggerida per la màquina i, addicionalment, poder-ho justificar⁷⁵¹. Aquest fenomen s'anomena “biaix

⁷⁴⁷Tribunal del Districte de La Haia, sentència de 5 de febrer de 2020, C/09/550982/ HA ZA 18-388, Case of SyRI, FJ 6.94.

⁷⁴⁸ALMADA, M., “Human intervention in automated decision-making...”, *op. cit.*, 2019, pàg. 7.

⁷⁴⁹*Idem*, pàg. 8.

⁷⁵⁰Tribunal del Districte de La Haia, sentència de 5 de febrer de 2020, C/09/550982/ HA ZA 18-388, Case of SyRI, FJ 6.101.

⁷⁵¹CASTETS-RENARD, C., “Accountability of Algorithms in the GDPR...”, *op. cit.*, 2019, pàg. 122.

d'automatització” (en anglès automation bias)⁷⁵². En aquest sentit, és interessant la reflexió de Huergo Lora en relació als algoritmes d'ajuda a les decisions administratives, on pot succeir que l'algoritme estigui mal dissenyat i calculi malament el resultat, que no permeti incorporar dades que siguin jurídicament rellevants en l'aplicació de la norma, o que la comoditat que proporciona indueixi a les persones usuàries a no comprovar la correcció del resultat, de forma que tot i que l'algoritme no prevalgui sobre la norma, i els resultats que no s'ajustin a la norma siguin il·legals, en molts casos prevalgui l'algoritme sobre la llei. Apunta que tot indica que en aquests casos el problema no és accedir al codi de programa, sinó aconseguir que s'apliqui la norma per sobre de l'error de l'algoritme⁷⁵³. Per aconseguir una interacció d'aquest tipus amb un sistema algorítmic cal analitzar els algoritmes com imaginariis sociotècnics en comptes de tecnologies discrecionals, permetent ampliar així el camp d'anàlisi⁷⁵⁴.

5.2.2. Dret de la persona a expressar el punt de vista

Tant el dret de la persona a expressar el punt de vista com el dret a oposar-se són una expressió del dret fonamental a la tutela judicial efectiva, reconeguts a l'article 47 de la CDFUE i l'article 13 del CEDH.

Com hem esmentat anteriorment, el dret a expressar el punt de vista requereix l'existència de la intervenció humana, perquè sinó, no hi hauria ningú a qui comunicar el punt de vista. Una vegada habilitat un canal per expressar el punt de vista, la persona podrà exposar aquells aspectes que en la seva opinió poden discernir de la decisió que ha pres el sistema algorítmic, aportant per exemple altres documents que no s'haguessin tingut en compte en el moment inicial o fent veure a la persona responsable del tractament que

⁷⁵² CITRON, D. K., “Technological Due Process”, *Washington University Law Review*, vol. 85, núm. 6, 2008, pàg. 1271 i 1272.

⁷⁵³HUERGO LORA, A., “Una aproximación a los algoritmos...”, *op. cit.*, 2020, pàg. 28.

⁷⁵⁴ GREEN, B., “The False Promise...”, *op. cit.*, 2020, pàg. 603; WIERINGA, M., ““Hey SyRI, tell me about algorithmic accountability”: Lessons from a landmark case”, *Data and Policy*, vol. 5, núm. 2, 2023, pàg. 3.

algunes dades de les que disposa i que han tingut rellevància en la decisió final són inexactes, incorrectes o incompletes per tal que es puguin rectificar⁷⁵⁵.

En primer lloc, l'existència pràctica d'un canal és tan important com que les persones que es veuen subjectes a una decisió automatitzada sàpiguen de l'existència d'aquest canal i de quin procediment han de seguir per poder expressar el seu punt de vista.

En segon lloc, per poder expressar un punt de vista s'ha de tenir una idea sobre la qüestió sobre la qual es vol donar l'opinió. Això sembla inversemblant en relació als sistemes algorítmics, ja que la persona hauria de saber concretament el funcionament del sistema, les dades que s'han emprat i quin valor se'ls ha donat. Per tenir una opinió ben formada s'haurien de destinar recursos per tal que les persones usuàries poguessin entendre com es prenen les decisions sobre elles. Com s'ha esmentat abans, exercir un dret sense tenir la informació necessària per fer-ho, inclús sense que la persona responsable del tractament la tingui (perquè l'algoritme pot funcionar de maneres imprevisibles), causa una sensació d'asimetria de poder i de manca de seguretat. En el context estatunidenc, es va establir en la sentència *Mathews contra Eldridge*⁷⁵⁶ del Tribunal Suprem un examen que aspira a proporcionar a la persona l'oportunitat de ser escoltada en un temps i d'una manera significativa però només si és rentable (cost-eficient). *Citron* estableix que en el context de l'anàlisi massiva de dades per tal que la persona sigui escoltada serà necessari l'accés al codi font o una audiència en relació amb la lògica de la decisió del programa d'ordinador, i que això sovint no serà rentable sota l'examen establert en *Mathews contra Eldridge*⁷⁵⁷, per tant lliga estretament l'oportunitat de la persona per ser escoltada amb l'accés a la informació.

En tercer lloc, quant a poder aportar informació que no s'hagi tingut en compte per poder rectificar la decisió, aquesta ha d'estar en un format que es pugui introduir en el sistema algorítmic. Aquesta és una qüestió rellevant, sobretot quan ens referim a la discriminació causada per decisions automatitzades, en el sentit que un dels problemes principals dels sistemes algorítmics és la seva incapacitat per tenir en compte o valorar aquella informació més subjectiva que no es troba estrictament en les dades.

⁷⁵⁵ PALMA ORTIGOSA, A., "Decisiones Automatizadas en el RGPD...", *op. cit.*, 2019, pàg. 33.

⁷⁵⁶ Tribunal Suprem dels Estats Units d'Amèrica del Nord, sentència de 24 de febrer de 1976, *Mathews v. Eldridge*, 424 U.S. 319, 333. Disponible a: <https://supreme.justia.com/cases/federal/us/424/319/> . Darrera visita el 2 de setembre de 2023.

⁷⁵⁷ CITRON, D. K., "Technological Due Process", *op. cit.*, 2008, pàg. 1284.

En quart lloc, en referència a la discriminació, aquesta es pot apreciar en comparació amb d'altres persones, i interessa disposar de dades comparatives per evidenciar-la. Com esmentàvem abans, és possible que ni les persones responsables del tractament puguin disposar d'aquesta informació, per consegüent, no sembla viable que la tinguin les persones subjectes de la presa de decisions.

Finalment, i en cinquè lloc, no s'especifica en l'RGPD de quina manera podrà la persona usuària expressar el seu punt de vista i quina repercussió podria tenir l'expressió del punt de vista en la decisió.

Se'ns plantegen aquí els mateixos conflictes que trobem quan parlem de la intervenció humana, per exemple, fins a quin punt amb la informació de la qual disposa una persona que no sap com funciona un sistema de presa de decisió, pot expressar un punt de vista suficientment motivat com per posar en dubte la decisió presa per una màquina amb la capacitat d'analitzar moltes més dades que ella. Hi ha aquí un clar desavantatge en el qual es veu atrapada la persona que es veu discriminada, ja que li serà molt difícil saber si la decisió està o no justificada en base als paràmetres emprats per l'algoritme. Esdevé també conflictiu què es farà en el cas que el punt de vista de la persona es consideri pertinent. Això suposarà que tindrà dret a ser objecte un altre cop de la presa de decisió, o que en el seu cas, la decisió la prendrà una persona? En base a quina informació podrà construir-se un punt de vista si no té pràcticament accés a cap informació llevat del resultat de la decisió?

5.2.3. Dret a impugnar la decisió

El dret a impugnar la decisió va unit al drets anteriorment analitzats, donat que davant la decisió presa per la màquina, la persona tindrà la possibilitat de dur a terme un tràmit d'audiència on podrà al·legar allò que estimi oportú (dret a expressar el seu punt de vista) a efectes de protegir els seus interessos, ocasionant que la persona responsable del tractament (dret a la intervenció humana) valori la petició i, en el cas pertinent, reconsideri la decisió presa per l'algoritme o incorpori noves dades al programa informàtic. No es tracta d'una acció processal davant l'ordre jurisdiccional competent, sinó de la possibilitat

que la persona interessada pugui recórrer la decisió davant la persona responsable del tractament en un context previ i addicional abans d'adreçar-se a l'autoritat competent o a la via judicial⁷⁵⁸. En aquest sentit, els articles 77 a 79 de l'RGPD també inclouen la possibilitat de plantejar queixes a l'autoritat supervisora competent, de desafiar la decisió de l'autoritat competent davant d'un tribunal i a obtenir solucions judicials davant les persones responsables del tractament.

Primerament, l'organització pública o privada que duu a terme aquesta gestió ha d'haver destinat els recursos humans i econòmics en crear aquest circuit i ha d'haver fet saber a les persones sobre les quals es pren una decisió de la seva existència i de la informació que poden tenir al seu abast per fer la impugnació. Segonament, en el cas de les causes de discriminació, com ja hem esmentat, l'opacitat i complexitat dels sistemes algorítmics i la falta d'informació en relació al seu funcionament dificulta l'exercici del dret, de la mateixa manera que impossibilita per part de la persona que vol impugnar la decisió d'aportar proves de la discriminació. Per tant, des de la perspectiva de la protecció contra la discriminació, aquest dret sembla poc efectiu. En el cas que la persona que veu tractades les seves dades consideri que ha patit una discriminació en la presa de decisió, no tindrà cap mena de prova per defensar-ho. Intuint que està essent discriminada, impugnarà la decisió i estarà, com ja ho està de partida, en una posició de desavantatge, ja que no només estarà completament subjugada a l'actuació de la persona responsable de tractament, sinó que inclús pot ésser que la persona responsable de tractament no sigui conscient que el sistema algorítmic està essent discriminatori.

En aquest sentit és evident que la posició del subjecte del tractament de dades enfront la persona responsable de dades és dèbil, podent ser comparada a la relació entre un consumidor i una empresa. La persona responsable del tractament realitzarà el tractament de les dades en el curs de les seves activitats professionals de manera estandarditzada i comptant amb els mitjans financers, la capacitat organitzacional i l'experiència tècnica⁷⁵⁹. Aquesta posició de desavantatge pot prevenir les persones que veuen les seves dades tractades d'exercir efectivament els seus drets⁷⁶⁰.

⁷⁵⁸PALMA ORTIGOSA, A., "Decisiones Automatizadas en el RGPD...", *op. cit.*, 2019, pàg. 34.

⁷⁵⁹WOLTERS, P.T.J., "The Control by and Rights of the Data Subject...", *op. cit.*, 2018, pàg. 7.

⁷⁶⁰BLUME, P., "The myths pertaining to the proposed General Data Protection Regulation", *International Data Privacy Law*, vol. 4, núm. 4, 2014, pàg. 270.

No es poden defensar els interessos propis si no es coneixen els arguments que es tenen en contra, de manera que difícilment es podran impugnar classificacions incorrectes, avaluacions basades en previsions imprecises o que afecten negativament les persones, sense entendre com s'ha pres la decisió i sota quines bases⁷⁶¹. Únicament es tindrà un autèntic dret a impugnar la decisió si la persona responsable del tractament informa específicament que no hi ha discriminacions directes, per exemple, basades en dades sensibles, i tampoc discriminacions indirectes, per exemple, aportant dades estadístiques que ho sostinguin. No obstant això, aquesta no és la pràctica que es duu a terme i el marc regulador actual manca de garanties en relació al dret potencial de reconsiderar la decisió, o, fins i tot, manca informació sobre com fer-ho. Què passa si l'individu impugna la decisió? Perdrà el dret a ser subjecte de la decisió si aquesta no fos automatitzada? Si realment s'obté una segona decisió, seria aquesta automatitzada o no automatitzada? No hi ha cap requeriment al respecte⁷⁶².

6. Conclusions parcials

En aquest capítol ens proposàvem respondre a la segona qüestió de recerca. Després d'abordar en el capítol anterior els punts febles de la normativa europea antidiscriminatòria, en aquest capítol hem abordat els punts febles que presenta el marc regulador europeu de protecció de dades per donar eines a les persones que s'han vist discriminades per desafiar la discriminació algorítmica.

Hem trobat diversos punts febles. En primer lloc, l'article 22 de l'RGPD només s'aplica quan no intervé cap persona en la decisió automatitzada, exclouent aquelles decisions en les quals ha intervingut una persona i han causat igualment discriminació, i presenta conceptes poc determinats com “efectes d'afectació significativament similar” o “salvaguardes mínimes” .

En segon lloc, l'RGPD no regula en l'articulat els biaixos algorítmics i la perspectiva adoptada en l'article 9 de l'RGPD de prohibir en el tractament de dades l'ús de certes

⁷⁶¹ROIG, A., *Las Garantías frente a las decisiones automatizadas. Del Reglamento General de Protección de Datos a la gobernanza algorítmica*, Barcelona, Bosch, 2020, pàg. 59.

⁷⁶²CASTETS-RENARD, C., “Accountability of Algorithms in the GDPR...”, *op. cit.*, 2019, pàg. 122.

dades personals, pot dificultar la detecció de biaixos, atesa la capacitat de l'anàlisi massiu de dades d'inferir dades personals a partir d'altres dades.

En tercer lloc, malgrat que els articles 12.1, 13.2 (f), 14.2 (g) i 15.1 (h) i el considerant 71 de l'RGPD proporcionen la possibilitat a les persones que es veuen subjectes a una decisió automatitzada d'obtenir informació sobre la presa de decisió, la informació que es dona és individual, pot interferir amb els drets de propietat intel·lectual, pot ser tan complexa que no sigui possible traduir-la en informació comprensible per a una persona no especialitzada, ja que pot ser intel·ligible, o no ser d'ajuda per proporcionar proves en un cas de discriminació algorítmica.

En quart lloc, trobem limitacions en les eines de gestió del risc que planteja l'RGPD -hem estudiat les avaluacions d'impacte de protecció de dades i auditories- per ser útils en la lluita enfront la discriminació algorítmica. Principalment, sembla difícil que es puguin aplicar a sistemes que evolucionen amb el temps, que puguin evitar que els paràmetres que les defineixen tinguin els mateixos biaixos (explícits i implícits) que es poden trobar en el tractament en si, i presenten el risc d'esdevenir processos merament formalistes i normatius sense més rellevància.

En cinquè lloc, la perspectiva individual de la protecció que planteja l'RGPD, confronta amb el factor col·lectiu no eliminable que trobem en la discriminació algorítmica. Els drets que confereix l'article 22.3 sempre es duren a terme individualment i no sembla probable que d'aquesta manera confereixin prou protecció enfront la discriminació algorítmica. A més, per provar discriminació algorítmica, el més probable és que sigui necessària informació sobre com el sistema algorítmic tracta a grups determinats de població.

Per tant, concloem, que la normativa de protecció de dades europea presenta punts febles que fan que per si sola pugui no donar suficients eines a les persones que s'han vist discriminades per desafiar la discriminació ocasionada a través de sistemes algorítmics. Tal com estableix Berendt, és obvi que de la discriminació ocasionada mitjançant sistemes algorítmics, no només són responsables els algorismes o les dades, de manera que no es pot confiar únicament en la protecció de dades per mitigar tots els efectes cumulats de

la discriminació⁷⁶³.

Si reprenem la segona qüestió de recerca que ens plantejàvem en aquest treball i que hem abordat tant en aquest capítol com en l'anterior, a saber, si el marc regulador europeu antidiscriminatori i el marc regulador europeu de protecció de dades són suficients per fer front a les discriminacions arran de les decisions automatitzades, considerem que no són suficients ni de manera dissociada ni de manera conjunta, ja que presenten punts febles de suficient entitat com per impossibilitar una protecció adequada. Arribats a aquesta conclusió, abordem en la segona part els punts forts d'aquestes normatives per veure fins a quin punt, sota certes condicions, poden ser útils per fer front a la discriminació algorítmica. També, donada la seva rellevància en el camp d'estudi, analitzem la PRIA.

⁷⁶³ BERENDT, B., "The AI Act Proposal: Towards the next transparency fallacy? Why AI regulation should be based on principles based on how algorithmic discrimination Works", a BMJV i ROSTALSKI, F. (eds.), *Künstliche Intelligenz - Wie gelingt eine vertrauenswürdige Verwendung in Deutschland und Europa?*, Tübingen, Mohr Siebeck, 2022, pàg. 42.

SEGONA PART. Potencialitats de la regulació europea per protegir a les persones de la discriminació causada per decisions automatitzades

En la segona part d'aquest treball, després d'haver estudiant en la primera part, per donar resposta a les dues primeres qüestions de recerca, com i per què les decisions automatitzades augmenten les possibilitats de discriminació, i com el marc regulador europeu antidiscriminació i de protecció de dades no són suficients per protegir a les persones de manera efectiva enfront la discriminació algorítmica, ens proposem donar resposta a la tercera i a la quarta qüestions de recerca, a saber, si algunes garanties de la normativa antidiscriminació i de protecció de dades podrien ser vàlides per protegir les persones de la discriminació algorítmica donades certes condicions, i si la PRIA ofereix noves garanties per protegir les persones o ajudar-les a desafiar decisions que presumptivament han produït discriminació algorítmica.

En el capítol IV, estudiem quins elements de la normativa europea antidiscriminació, sota certes condicions, poden ser útils per a la regulació de la discriminació algorítmica; en el capítol V, analitzem quins elements de la normativa de protecció de dades, sota certes condicions, poden ser útils per a la regulació de la discriminació algorítmica; i finalment, en el capítol VI, examinem la PRIA i les garanties que pot aportar a la regulació de la discriminació algorítmica.

Capítol IV. Possibilitats de la regulació antidiscriminació per abordar la discriminació algorítmica

En aquest capítol i en el següent abordem la tercera qüestió de recerca, a saber, si donades certes condicions, algunes garanties de la regulació antidiscriminació i de protecció de dades podrien ser vàlides per protegir les persones de la discriminació algorítmica. Per respondre a aquesta pregunta, en aquest capítol IV del treball presentem els punts forts que presenta el marc regulador europeu antidiscriminatori i si aquests, sota certes condicions, poden mitigar els punts febles que hem plantejat en el capítol II o ser directament útils per a l'abordatge de la discriminació algorítmica. Com a mesures de mitigació s'ha plantejat la idoneïtat de llistes obertes d'atributs protegits i l'homogeneïtzació de la protecció que confereix la legislació derivada de la UE en relació amb tots els atributs protegits (1), i la promoció de l'ús d'una perspectiva estructural i interseccional per abordar millor la discriminació que poden causar els sistemes algorítmics (2). Com a elements útils ja existents en la normativa, estudiem com el TJUE ha flexibilitzat a través de la seva jurisprudència la concepció de la discriminació directa i indirecta (3), i els elements de la jurisprudència del TJUE que es podrien adaptar i que facilitarien la prova de discriminació algorítmica (4). Acabem el capítol amb un apartat de conclusions parcials (5).

1. Llistes obertes d'atributs protegits i homogeneïtzació de la protecció

En el primer apartat del capítol II, hem abordat els punts febles de la regulació antidiscriminació de la UE per fer front a la discriminació algorítmica, que estan relacionats, per una banda, amb la concreció en l'article 19 del TFUE d'una llista exhaustiva dels atributs pels quals s'habilita al Consell i al Parlament Europeu a dur a terme accions per combatre la discriminació, i per l'altra banda, amb que la protecció del dret derivat no és igual per a tots els atributs protegits. En aquest apartat analitzem les solucions que s'han proposat per millorar la protecció. En primer lloc, l'ampliació de la protecció a més atributs (1.1) i en segon lloc, l'homogeneïtzació de la protecció envers la

discriminació que es brinda en el dret derivat perquè sigui igual per tots els atributs protegits (1.2).

1.1. Llistes obertes d'atributs protegits

Tal com indicàvem en el l'apartat 1.2 del capítol II, s'ha fet palès que la redacció de l'article 19 del TFUE⁷⁶⁴, que estableix una llista exhaustiva d'atributs sobre els quals les institucions de la UE poden dur a terme accions per combatre la discriminació (sexe, origen racial o ètnic, religió o les conviccions, discapacitat, edat o orientació sexual), sense que sigui possible una ampliació per analogia a altres motius no expressament enumerats⁷⁶⁵, pot evidenciar limitacions per abordar la discriminació algorítmica⁷⁶⁶.

S'ha reiterat, per part del TJUE, en relació amb la Directiva 2000/78/CE (que dimana de l'article 13 CE, actualment article 19 TFUE)⁷⁶⁷, que no procedeix ampliar el seu àmbit d'aplicació per analogia a altres tipologies de discriminació, a més d'aquelles basades en els motius enumerats amb caràcter exhaustiu en l'article primer de la pròpia Directiva⁷⁶⁸, és a dir, que no es poden afegir motius en relació als quals la Directiva 2000/78/CE prohibeix tota discriminació (discapacitat, orientació sexual, religió o creences o edat). Entenem que el mateix criteri pot aplicar-se a les altres directives antidiscriminació⁷⁶⁹.

⁷⁶⁴ L'article 19 del TFUE.

⁷⁶⁵ REQUENA CASANOVA, M., "La tutela judicial del principio general de igualdad de trato en la Unión Europea: una jurisprudencia expansiva basada en una jerarquía de motivos discriminatorios", *Revista de Derecho Comunitario Europeo*, núm. 40, pàg. 771.

⁷⁶⁶ GERARDS, J. i XENIDIS, R., *Algorithmic Discrimination in Europe...*, op. cit., 2021, pàg. 49.

⁷⁶⁷ RIVAS VALLEJO, P., "Un matiz nuevo en la delimitación del ámbito subjetivo de protección de las directivas antidiscriminación: diferencia de trato entre sujetos de la misma categoría protegida", *Revista de Jurisprudencia Laboral*, núm. 1/2021, 2021, pàg. 6.

⁷⁶⁸ TJCE (Gran Sala), sentència d'11 de juliol de 2006, C-13/05, Sonia Chacón Navas contra Eurest Colectividades SA, par. 56; TJUE (Sala Quarta), sentència de 18 de desembre de 2014, C-354/13, Fag og Arbejde (FOA) contra Kommunernes Landsforening (KL), par. 36; TJCE (Gran Sala), sentència de 17 de juliol de 2008, C-303/06, S. Coleman contra Attridge Law and Steve Law, par. 46; TJUE (Sala Segona), sentència de 9 de març de 2017, C-406/15, Petya Milkova contra Izpalnitelen direktor na Agentsiata za privatizatsia i sledprivatizatsionen kontrol, amb la intervenció de Varhovna administrativna prokuratura, par. 34; TJUE (Sala Vuitena), sentència de 8 d'octubre de 2020, C-644/19, FT contra Universitatea «Lucian Blaga» Sibiu, GS y otros, HS, Ministerul Educației Naționale, par. 31.

⁷⁶⁹ TJUE (Sala Primera), sentència de 6 d'abril de 2017, C-668/15, Jyske Finans A/S contra Ligebehandlingsnavnet, par. 24. Es descarta que l'article 3.2 i el considerant 13 de la Directiva 2000/43/CE afecti a les diferències de tracte per motiu de nacionalitat.

No obstant això, com els atributs no estan definits sota la regulació europea, els tribunals nacionals són responsables de dibuixar contextualment els seus límits, un exercici que s'hauria de dur a terme amb la finalitat de assegurar l'efectivitat de la regulació⁷⁷⁰. Per exemple, les qüestions relatives a discernir si una conducta constitueix o no assetjament⁷⁷¹ se solen determinar en l'àmbit nacional, abans de remetre els assumptes al TJUE⁷⁷². La complexitat inherent, natura composta i heterogeneïtat dels atributs protegits s'hauria de reconèixer per part dels tribunals⁷⁷³, que han arribat a admetre la interpretació abstracta del motiu concret⁷⁷⁴. Igualment, s'ha de tenir en compte que les proteccions del dret antidiscriminatori dels contextos de tractament de dades contemporanis i futurs tenen aplicació limitada, en haver-hi dificultats per identificar tractament diferencial en base a atributs protegits, especialment quan aquests són abstractes o interseccionals⁷⁷⁵.

L'elaboració de perfils complica en dues vessants la noció de resultat discriminatori en relació amb un atribut protegit: primerament, perquè es poden utilitzar indicadors i, segonament, perquè es poden utilitzar categories que no tenen un significat clar per la interpretació humana. Això suposa que la discriminació no concernirà a cap atribut protegit sinó que concernirà factors com el salari, el codi postal, el cotxe o combinacions algorítmiques complexes de diversos atributs⁷⁷⁶. Mann i Matzner defensen que pel que fa a la naturalesa abstracta de l'elaboració de perfils i les inferències, pot ser que no sigui possible identificar els motius de discriminació segons els motius protegits específics (una tipologia de discriminació molt difícil de detectar ja que no provoca formes d'estigmatització que es reconeguin socialment), i que una perspectiva més diversificada de la lluita envers la discriminació pot ser una via a explorar, a fi que combinacions i correlacions algorítmiques complexes també puguin ser reconegudes com a

⁷⁷⁰ XENIDIS, R., "Turning EU equality law to algorithmic discrimination...", *op. cit.*, 2020, pàg. 750.

⁷⁷¹ S'especifica en l'article 2.2 de la Directiva 2006/54/CE que el concepte de discriminació inclourà l'assetjament, l'assetjament sexual i l'ordre de discriminar.

⁷⁷² ARCOS VARGAS, M., "La no discriminación en el derecho derivado...", *op. cit.*, 2022, pàg. 25. Que els tribunals nacionals puguin dibuixar contextualment els límits dels atributs protegits també pot distorsionar la uniformitat de l'aplicació del dret de la UE, donat que, per exemple, existeixen diferents percepcions del que suposa assetjament en funció de l'entorn social on es produeixi.

⁷⁷³ XENIDIS, R., "Turning EU equality law to algorithmic discrimination...", *op. cit.*, 2020, pàg. 750.

⁷⁷⁴ ARCOS VARGAS, M., "La no discriminación en el derecho derivado...", *op. cit.*, 2022, pàg. 24.

⁷⁷⁵ MANN, M. i MATZNER, T., "Challenging algorithmic profiling...", *op. cit.*, 2019, pàg. 3. Mann i Matzner analitzen l'elaboració de perfils algorítmica com un procés de construcció de coneixement basat en grans bases de dades que sovint no tenen cap relació directa amb els atributs protegits de la regulació antidiscriminació, podent formar categories no representatives, interseccionals i complexes que suposin desavantatges sistèmics per a grups de persones (que ho desconeixen).

⁷⁷⁶ *Ídem*, pàg. 4.

discriminatòries⁷⁷⁷.

En aquest sentit, Gerards i Zuiderveen fan una anàlisi dels sistemes existents per abordar la no discriminació i la igualtat de tractament. Determinen l'existència de sistemes completament oberts (per exemple, l'article 20 de la CDFUE que estableix que "*tota persona és igual davant la llei*", sense enumerar atributs protegits); sistemes completament tancats (per exemple, l'article 19 del TFUE que habilita a la Unió Europea per la creació de les directives antidiscriminació i que només enumera sis atributs protegits); i sistemes híbrids (per exemple, els articles 14 del CEDH i 21 de la CDFUE en què s'enumeren una llista d'atributs protegits i alhora s'afegeixen fórmules com "*o qualsevol altre situació*" o "*es prohibeix tota discriminació, i en particular...*", respectivament)⁷⁷⁸. En relació amb les justificacions de l'existència de discriminació, constaten el caràcter híbrid del mecanisme actual, és a dir, un sistema de justificacions tancat per la discriminació directa -malgrat a la pràctica poden deixar marge de maniobra: en l'article es posa l'exemple de justificacions de discriminació per raó d'edat en la Directiva 2000/78/CE- i un sistema obert de justificacions per la discriminació indirecta⁷⁷⁹.

Consideren Gerards i Zuiderveen que el sistema més adequat per abordar les especificitats de la discriminació algorítmica és un sistema híbrid en relació amb la llista d'atributs protegits i un sistema obert de justificacions. Argumenten que un sistema híbrid d'atributs protegits permet exemplificar els atributs que es considera que requereixen especial protecció mitjançant processos legislatius, donant la possibilitat a organitzacions i tribunals de comparar-los amb altres atributs o indicadors que siguin similars perquè reflecteixin biaixos, prejudicis o estereotips de manera comparable⁷⁸⁰. També, com les decisions preses mitjançant sistemes algorítmics poden crear noves formes de desavantatge que es basen en altres característiques no especialment esmentades (p. e. estatus socioeconòmic o pobresa), i les llistes no exhaustives permeten desafiar casos concrets de desavantatge no basats en característiques especialment enumerades, s'adapten millor a les particularitats de la diferenciació mitjançant sistemes

⁷⁷⁷ *Ídem*, pàg. 3, 4 i 6. Consideren també que la lluita envers les noves formes de discriminació a través de les proteccions antidiscriminatòries tenen el risc de continuar les pràctiques colonials de generar salvaguardes creant exclusions.

⁷⁷⁸ GERARDS, J. i ZUIDERVEEN, F., "Protected Grounds and the System...", *op. cit.*, 2022, pàg. 22-28.

⁷⁷⁹ *Ídem, ibídem*.

⁷⁸⁰ *Ídem*, pàg. 50-51.

d'intel·ligència artificial. Conclouen que les directives antidiscriminació europees no són ideals com a base legal per fer front a la discriminació algorítmica i, en canvi, tant l'article 14 de la CEDH com l'article 21 de la CDFUE suposen provisions útils en relació amb la discriminació algorítmica⁷⁸¹.

En el cas de les justificacions, consideren que és difícil establir una llista exhaustiva de justificacions de tractament desigual que sigui capaç de fer front als reptes específics de la diferenciació causada pels algorismes. En vista de l'espectre de situacions en les quals pot ocórrer la discriminació algorítmica i la diversitat de diferenciació que la intel·ligència artificial pot causar, pot no ser possible arribar a un conjunt de justificacions integral, útil i que s'entengui, que es pugui incloure fàcilment en el dret antidiscriminatori. Gerards i Zuiderveen defensen, per consegüent, que únicament un sistema obert de justificacions és possible, de manera que per totes les possibilitats de diferenciació algorítmica, les organitzacions que emprin intel·ligència artificial han de poder mostrar una justificació raonable⁷⁸². Des de la nostra perspectiva, no hauria de ser possible un sistema de justificacions, de manera que hauria de prohibir-se l'ús de qualsevol sistema algorítmic que es determinés que és discriminatori.

En definitiva, part de la doctrina ha proposat i considera necessàries lògiques i proteccions més àmplies, que operin amb independència d'atributs protegits específics, per exemple abraçant la redacció de provisions com l'article 14 del CEDH que prohibeix la discriminació en base a qualsevol atribut⁷⁸³, o l'article 21 de la CDFUE, que prohibeix

⁷⁸¹ *Ídem*, pàg. 52.

⁷⁸² *Ídem*, pàg. 53-54.

⁷⁸³ MANN, M. i MATZNER, T., "Challenging algorithmic profiling...", *op. cit.*, 2019, pàg. 6; ARCOS VARGAS, M., "La no discriminación en el derecho derivado...", *op. cit.*, 2022, pàg. 32 i 39. L'autora fa certes especificacions per l'aplicació del marc normatiu del CEDH. Malgrat que la seva llista d'atributs protegits sigui més oberta, es pot aplicar l'article 14 sempre que estigui afectat un dels drets recollits en el conveni, o l'estat membre sigui signatari del Protocol 12 al CEDH (que només ha estat ratificat per 17 dels 47 estats membres del Consell d'Europa, entre ells 6 membres de la UE), que prohibeix de la discriminació a la protecció de tots els drets reconeguts en l'àmbit nacional encara que no estiguin inclosos en l'àmbit dels drets emparats pel CEDH. Això suposa que els estats membres de la UE tenen diferents obligacions derivades de la legislació europea contra la discriminació i hi ha àrees d'aplicació del CEDH que queden fora de l'àmbit de la normativa de la UE. No només això, sinó que també hi ha diferències rellevants en relació amb les persones destinatàries de la protecció derivada de la legislació de la UE i el CEDH. El CEDH protegeix a totes les persones situades sota la jurisdicció d'un estat membre, siguin o no ciutadans/es d'aquest i es trobin o no en el seu territori nacional, sempre que es trobin en àrees situades sota el control efectiu de l'Estat (com els territoris ocupats). En canvi, les directives antidiscriminació contenen diverses exclusions de l'aplicació a nacionals de tercers països, és a dir, ciutadans/es d'Estats que no siguin membres de la UE.

tota discriminació⁷⁸⁴. La referència en aquestes normes a altres motius o les formules com “qualsevol altre situació” o “tota discriminació” deixa oberta la possibilitat que altres atributs emprats en relació amb les pràctiques d’elaboració de perfils puguin ser considerats discriminatoris⁷⁸⁵.

No obstant això, cal fer una última puntualització. Wachter exposa que amb el funcionament de sistemes algorítmics, s’eixampla el gruix de víctimes de discriminació. Grups de persones percebudes com a similars entre si poden ser tractades injustament (per exemple, oferint-los préstecs amb interessos més alts o productes financers de més risc), sense ser singularitzades sobre la base d’atributs protegits. D’aquesta manera, poden patir tracte menys favorable “persones que tenen un gos”, “persones que van néixer un dimarts” o “persones que juguen a videojocs”⁷⁸⁶. Des de la perspectiva antilibertatària, aquestes persones no haurien de ser destinatàries de la protecció antidiscriminació, atès que no viuen una situació de subordinació o exclusió, sinó que únicament viuen una situació de tracte desigual sense cap relació amb una història d’opressió. En aquest sentit, el que Wachter considera discriminatori, per a nosaltres, en no suposar una perpetuació del tracte desfavorable a grups de persones històricament oprimides, és un tracte desigual. En el cas que “persones que tenen un gos” es relacioni amb un grup considerat històricament oprimat, llavors sí que es podria considerar discriminació. Si no es així, simplement no se’ls ha tractat de la mateixa manera, sense una justificació lògica per fer-ho -això també suposa un problema però no del caire relatiu a la discriminació de grups històricament oprimits. Això és degut a que l’entitat d’un tracte desigual per tenir un gos no té les mateixes implicacions socials que un tracte desigual per ser dona (com també, malgrat en diferents graus, els altres atributs protegits per la legislació). Un tracte desigual per ser dona implica unes connotacions d’opressió històriques i culturals que no poden equiparar-se a tenir un gos.

És cert que la línia de connexió sobre la qual es basa la tutela antidiscriminatòria té un fonament necessàriament axiològic, és a dir, parteix d’uns valors determinats⁷⁸⁷, fet que afecta marcadament les prioritats de la legislació antidiscriminació, i per consegüent, els

⁷⁸⁴ GERARDS, J. i XENIDIS, R., *Algorithmic Discrimination in Europe...*, *op. cit.*, 2021, pàg. 9; XENIDIS, R., “Turning EU equality law to algorithmic discrimination...”, *op. cit.*, 2020, pàg. 755.

⁷⁸⁵ SCHREURS, W., “Cogitas, Ergo Sum. The role of data protection...”, *op. cit.*, 2008, pàg. 260.

⁷⁸⁶ WACHTER, S., “Affinity profiling and discrimination...”, *op. cit.*, 2020, pàg. 415.

⁷⁸⁷ RODRÍGUEZ-PIÑERO y BRAVO-FERRER, M., “Los contornos de la discriminación”, *op. cit.*, 2022, pàg. 18.

atributs que es considera que necessiten especial protecció sorgeixen en base a un consens de l'existència de raons històriques d'opressió que causen desigualtats patents en la societat. Així i tot, el fet que un llistat d'atributs a protegir de discriminació sigui obert suposa que no necessàriament s'aplica únicament a grups socials que han estat perseguits en el passat. Les noves causes de discriminació poden reconèixer-se com a conseqüència d'una presa de consciència respecte a sistemes d'opressió històricament arrelats, com ha succeït amb l'orientació sexual, però també poden aparèixer de manera sobrevinguda en relació amb canvis econòmics, polítics, tecnològics, culturals o socials, com succeeix amb la discriminació pel contingut del material genètic. Així doncs, no només es condemnen determinades situacions presents en la realitat històrica o actual, sinó que també una sèrie de conductes que responen a una lògica de persecució o segregació i que, en conseqüència, poden generar en un futur nous desequilibris socials que afectin a la dignitat de la persona⁷⁸⁸. L'obertura de la llista d'atributs protegits l'entendem en el sentit que no es limitin a protegir únicament l'atribut esmentat sinó que també protegeixin aquells indicadors que a l'emprar-se tenen les mateixes conseqüències que si s'hagués utilitzat l'atribut protegit, o permetin incloure atributs que puguin generar, ja en el present (estatus socioeconòmic) o en el futur (material genètic) desequilibris socials. No entenem que aquesta llista pugui incloure altres atributs que arran dels sistemes algorítmics puguin causar tracte desigual, si aquests no estan relacionats amb sistemes d'opressió i dominació.

De l'article 19 del TFUE deriven les directives antidiscriminació en base als atributs protegits que s'enumeren. Aquestes directives no creen un règim de protecció igual, sinó que segons l'atribut protegit el règim de protecció és diferent. Analitzem, tot seguit, quines propostes s'han fet per mitigar aquesta diferència de protecció en l'àmbit de la discriminació algorítmica.

1.2. Homogeneïtzació de la protecció

⁷⁸⁸ ÁLVAREZ DEL CUVILLO, A., "El problema de la discriminación inversa...", *op. cit.*, 2022, pàg. 193 i 194.

Com ja hem esmentat, les directives antidiscriminació estableixen un règim més complet en els àmbits en els quals es considera que la igualtat i la no discriminació mereixen especial protecció, definint en relació a aquests de manera més extensa les categories protegides i les classes de discriminació, constituint un mínim que els estats membres han d'incorporar en les seves legislacions internes⁷⁸⁹. D'aquesta manera, la protecció en les directives antidiscriminació cobreix tres àrees: treball, sistema de benestar i accés a béns i serveis. La Directiva 2000/43/CE és l'única que s'aplica en les tres àrees, la Directiva 2000/78/CE només s'aplica en l'àmbit laboral, la Directiva 2006/54/CE i la Directiva 2004/113/CE s'apliquen en els àmbits de treball i l'accés a béns i serveis però no a l'accés al sistema de benestar⁷⁹⁰.

Justament perquè en certs àmbits la protecció està més desenvolupada, s'evidencia que en d'altres no existeix, generant importants limitacions, com per exemple, la prohibició de discriminació en l'accés a béns i serveis aplicada únicament a l'ètnia i el sexe però no a les altres categories protegides; la no inclusió de la publicitat en l'àmbit de la Directiva 2004/113/CE; o que la protecció enfront la discriminació per raó de religió, conviccions, edat, orientació sexual i discapacitat només es tingui en compte en l'àmbit del treball⁷⁹¹. Per consegüent, i com ja hem abordat en l'apartat 1.3 del capítol II, la legislació antidiscriminatòria a la UE té abast asimètric⁷⁹². Tanmateix, molts estats membres estableixen un conjunt coherent de drets per a cada una de les característiques, tant en l'àmbit de l'ocupació com en el de l'accés a béns i serveis⁷⁹³.

Respecte a les directives s'han plantejat diverses millores perquè siguin més efectives en relació amb la discriminació algorítmica. Per exemple, en el cas de la discriminació algorítmica per raó de raça o origen ètnic, estereotips i prejudicis poden impregnar els algorismes emprats per determinar la distribució d'anuncis i, en conseqüència, l'accés a béns i serveis. Malgrat que la igualtat en relació a l'ètnia està protegida a la UE en relació als béns i serveis, en la Directiva 2000/43/CE no es fa cap menció sobre els mitjans de

⁷⁸⁹Aquestes directives només estableixen uns requisits mínims i un marc regulador general que necessita ser transposat en la llei nacional pels estats membres. Molts estableixen una protecció més exhaustiva que aquella proposada per les directives, deixant la UE amb diferents estàndards segons l'estat membre.

⁷⁹⁰ARCOS VARGAS, M., "La no discriminación en el derecho derivado...", *op. cit.*, 2022, pàg. 33.

⁷⁹¹SORIANO ARNANZ, A., "Decisiones automatizadas y discriminación...", *op. cit.*, 2021, pàg. 30.

⁷⁹²GELLERT, R. *et al.*, "A comparative analysis...", *op. cit.*, 2013, pàg. 63-66. En l'article s'especifica que un tractament diferencial pot tenir moltes formes i materialitzar-se en qualsevol tipus d'acció, per això la legislació antidiscriminació no es limita a un cert tipus de comportament o pràctica, posant la seva atenció més en el resultat que no en el procés.

⁷⁹³ALLEN, R. i MASTERS, D., "Artificial Intelligence: the right to...", *op. cit.*, 2020, pàg. 591.

comunicació ni la publicitat. Tampoc hi ha cap excepció explícita com en la Directiva 2004/113/CE, per tant seria lògic assumir que aquests camps estan coberts. No obstant això, la inclusió dels mitjans de comunicació i la publicitat com a àmbits protegits per la directiva 2000/43/CE seria interessant, donada la rellevància que tenen quan ens referim a la discriminació algorítmica⁷⁹⁴.

També, les legislacions nacionals promulgades arran d'aquestes directives poden tenir un àmbit material més ampli en escollir no tenir en compte les excepcions disponibles (per exemple, estendre l'aplicació de les disposicions de la Directiva 2004/113/CE als mitjans de comunicació, a la publicitat i a l'educació), i per tant, ofereixen més possibilitats, en aquesta qüestió, de prevenir i redreçar la discriminació algorítmica.

El dret derivat europeu té en compte la majoria d'atributs protegits (malgrat s'atorguin graus de protecció diferents segons l'atribut)⁷⁹⁵ en l'àmbit laboral, mentre que els algoritmes s'empren en altres àrees com la distribució de beneficis socials i les taxes, la salut i les assegurances, entre d'altres. El conjunt normatiu de dret derivat, hauria de completar-se amb una norma que permetés aconseguir la protecció integral contra les discriminacions per aquests motius. En aquest sentit es va presentar la Propuesta de Directiva del Consejo por la que se aplica el principio de igualdad de trato entre las personas independientemente de su religión o convicciones, discapacidad, edad u orientación sexual⁷⁹⁶ per part de la Comissió Europea l'any 2008, amb la finalitat d'aplicar el principi d'igualtat de tracte fora del context laboral⁷⁹⁷, però actualment encara no s'ha arribat a cap acord.

En resum, s'ha plantejat la necessitat de proporcionar un esquema de protecció més ampli, que operi amb independència d'atributs protegits específics, abraçant la redacció de provisions com l'article 14 del CEDH que prohibeix la discriminació en base a qualsevol atribut, o l'article 21 de la CDFUE, que prohibeix tota discriminació. En aquestes provisions, s'han d'interpretar, llavors, els atributs enumerats com a exemples, donant la possibilitat a organitzacions i tribunals de comparar-los amb altres atributs o indicadors que siguin similars perquè reflecteixin biaixos, prejudicis o estereotips de manera

⁷⁹⁴GERARDS, J. i XENIDIS, R., *Algorithmic Discrimination in Europe...*, *op. cit.*, 2021, pàg. 60.

⁷⁹⁵SORIANO ARNANZ, A., "Discriminación algorítmica: garantías...", *op. cit.*, 2022, pàg. 163.

⁷⁹⁶COMISIÓN DE LAS COMUNIDADES EUROPEAS, *Propuesta de Directiva del Consejo por la que se aplica el principio de igualdad de trato entre las personas independientemente de su religión o convicciones, discapacidad, edad u orientación sexual*, COM (2008) 426-final, Bruselas, 2 de julio de 2008.

⁷⁹⁷ARCOS VARGAS, M., "La no discriminación en el derecho derivado...", *op. cit.*, 2022, pàg. 21.

comparable. Quant a les directives antidiscriminació, s'ha subratllat la necessitat d'equiparar la protecció que es brinda a tots els atributs protegits en relació als àmbits materials (treball, accés a béns i serveis, i sistema del benestar). Tot seguit estudiem les millores que pot aportar el reconeixement de la incidència de la discriminació estructural i interseccional en la discriminació causada mitjançant sistemes algorítmics.

2. Reconeixement del caràcter grupal en la discriminació algorítmica

El dret antidiscriminatori, actualment, reconeix fins a cert punt la dimensió col·lectiva en establir grups de població especialment protegits a través de categories, però no resol les tensions que la perspectiva de grup genera en relació amb la tradició jurídica basada en drets individuals⁷⁹⁸.

Com hem esmentat en l'apartat 2 del capítol II, el dret antidiscriminatori entès des de la perspectiva antisubordinatòria té la intenció de donar respostes no únicament a demandes individuals (inclús grupals) de discriminació, sinó que es configuri un entramat normatiu amb l'objectiu de transformar les estructures de desigualtat (incloent noves categories com la discriminació per associació, múltiple o interseccional)⁷⁹⁹, com també el reconeixement de la discriminació estructural. També, tracta de subministrar arguments orientats a establir límits a determinats desenvolupaments de la intel·ligència artificial que erosionen els drets humans, que provoquen un increment de les desigualtats o que exacerbem les discriminacions existents⁸⁰⁰.

Per identificar una situació de desavantatge de les persones i grups que reben tracte discriminatori des d'una perspectiva antisubordinatòria, els conceptes com discriminació directa i indirecta són insuficients per determinar el patró de discriminació estructural, prejudicis socials i estereotips que conformen la norma que s'ha de jutjar discriminatòria o no, de manera que, segons la doctrina crítica amb l'enfocament del dret antidiscriminatori "anticlassificació", cal incorporar criteris interpretatius auxiliars que es

⁷⁹⁸ SEBASTIANI, L., "Investigando los límites de la lucha...", *op. cit.*, 2021, pàg. 757. L'autor, en relació amb el racisme, detecta que des de la perspectiva legalista, se subestima el component històric del racisme i es desdibuixa el reconeixement de la seva dimensió col·lectiva.

⁷⁹⁹ AÑÓN ROIG, M.J., "Desigualdades algorítmicas...", *op. cit.*, 2022, pàg. 42.

⁸⁰⁰ *Ídem*, pàg. 46.

trobin en l'enfocament de la discriminació estructural (2.1) i en la perspectiva de la interseccionalitat⁸⁰¹ (2.2).

2.1. Reconeixement de la discriminació estructural

La desigualtat estructural situa l'estat de privació en què es troben la majoria de grups oprimits, en els processos socials i en les pràctiques culturals que defineixen el seu estatus, en comptes d'entendre'l com a conseqüència de desavantatges transitoris i fortuïts producte de la mala sort o de decisions individuals desencertades. En aquest sentit, el criteri de la discriminació estructural pot ser entès com un criteri interpretatiu de suport a la discriminació indirecta⁸⁰², i exigeix al dret antidiscriminatori el reconeixement dels sistemes d'opressió⁸⁰³.

Els trets característics de la discriminació estructural mostren que es tracta de processos socials difusos, sistèmics -al marge de la intencionalitat o voluntat de les persones individualment considerades-, que es reproduïxen institucionalment ja que travessen i projecten totes les dimensions de l'existència, tant en l'àmbit públic i social, com en l'àmbit domèstic, encreuant-se al seu torn amb les variables socials de major rellevància, ocasionant una incidència profunda que afecta la capacitat mateixa de presa de decisions i la formació de preferències⁸⁰⁴.

Tanmateix, com hem expressat en l'apartat 2 del capítol II, el problema de les relacions de domini/subordinació es reconduïx al gaudi de drets individuals (quan no hi ha un gaudi complet dels drets individuals, hi ha discriminació, en tant que això representa la ruptura de la regla d'igualtat estructural en una societat en la qual el dret presumeix la igualtat de totes les persones en relació amb el seu gaudi). Però abordar la major dificultat de gaudi dels drets individuals, com a conseqüència de les relacions de domini/subordinació, no és el mateix que abordar les relacions de domini/subordinació

⁸⁰¹ AÑÓN ROIG, M. J., "Principio Antidiscriminatorio...", *op. cit.*, 2013, pàg. 146 i 147.

⁸⁰² *Ídem*, pàg. 150.

⁸⁰³ *Ídem*, pàg. 147.

⁸⁰⁴ AÑÓN ROIG, M. J., "Principio Antidiscriminatorio...", *op. cit.*, 2013, pàg. 148. L'autora apunta que desconèixer o ocultar els patrons de discriminació estructural pot portar a presentar com una decisió lliure una cosa que, en realitat, no és sinó una preferència adaptativa o una decisió marcada per una necessitat, invalidant la presumpció de consentir de manera lliure i informada.

en si⁸⁰⁵. Una acció discriminatòria no incideix només en el gaudi dels drets individuals sinó que és una expressió del domini que trenca la regla (de presumpció) d'igualtat de poder entre grups socials⁸⁰⁶. És a dir, la idea subjacent al concepte de discriminació estructural és trencar amb la presumpció segons la qual pel dret no hi ha discriminació mentre no existeixi una conducta imputable individualment⁸⁰⁷.

Llavors, què suposaria el reconeixement de la discriminació estructural per a l'abordatge de la discriminació algorítmica? En primer lloc, la consideració que el tracte que produeixi una discriminació, encara que es manifesti de forma individualitzada, sigui interpretat a la llum de la posició social estructural de subordinació, és a dir, no només reconèixer que la vida en societat genera grups discriminats i que el dret ha d'intervenir en la discriminació, sinó també en admetre que no hi ha dret socialment neutre ni neutral, i que, si hi ha discriminació és perquè ja sigui per acció o omissió, el dret també discrimina⁸⁰⁸.

En segon lloc, l'admissió que els algoritmes repliquen i perpetuen les estructures de dominació i tenen efectes negatius pels grups discriminats i per les seves possibilitats de participació en la societat, en facilitar la identificació i la singularització de les persones, grups i segments socials sobre els quals recau la discriminació⁸⁰⁹. Normalment, s'atribueix com a causa dels resultats discriminatoris del sistema un perjudici irracional o el caràcter erroni o esbiaixat de les dades seleccionades, quan en la majoria dels casos són un reflex de desigualtats estructurals i el seu reconeixement deriva del principi antisubordinació⁸¹⁰.

En tercer lloc, facilitar un abordatge de la discriminació que com a element central inclou el perjudici cap al grup que comparteix l'atribut protegit. En les entrevistes fetes per Gerards i Xenidis, es destaca per part de persones expertes en discriminació de França, Alemanya i Polònia, que un dels problemes inherents de la legislació antidiscriminatòria és que es focalitza molt en els casos individuals (on s'obliga a presentar un cas de

⁸⁰⁵ BARRÈRE, M.A., "Iusfeminismo y derecho antidiscriminatorio...", *op. cit.*, 2008, pàg. 62-66.

⁸⁰⁶ *Ídem*, pàg. 63.

⁸⁰⁷ *Ídem*, pàg. 56.

⁸⁰⁸ *Ídem*, pàg. 67-68; SEBASTIANI, L., "Investigando los límites de la lucha...", *op. cit.*, 2021, pàg. 760. L'autor exposa que és necessària una desmitificació de la presumpta "innocència" del dret respecte al racisme.

⁸⁰⁹ AÑÓN ROIG, M.J., "Desigualdades algorítmicas...", *op. cit.*, 2022, pàg. 35.

⁸¹⁰ *Ídem*, pàg. 45 i 46.

discriminació indirecta envers un grup ben definit) quan moltes instàncies de la discriminació algorítmica deriven de problemes estructurals⁸¹¹. Per consegüent, donat l'increment de l'ús d'algoritmes en la presa de decisions en totes les àrees de la vida, suposa el risc d'amplificar i magnificar les desigualtats estructurals existents i, per tant, la perspectiva del dret antidiscriminatori basada en l'abordatge individual es mostra insuficient (en assumir-se que es tracta d'un problema puntual on l'escenari clau és la litigació i la finalitat merament reparadora).

En vista de les anteriors consideracions, el dret antidiscriminació hauria de reclamar que els models algorítmics comptessin amb garanties, prohibicions clares de discriminació (en les dades i en les classificacions) i mecanismes que possibilitessin l'examen, el coneixement i el control de les diferents formes de discriminació, així com articular respostes sobre aquestes qüestions. També s'hauria d'ampliar l'impacte de la seva acció, considerant que la seva finalitat rau en la transformació de les estructures socials creadores dels sistemes d'opressió i adoptant una visió prioritàriament centrada en les causes de la discriminació estructural i en la lògica subjacent a l'ús de la intel·ligència artificial⁸¹².

A continuació estudiem la discriminació interseccional, que pretén fer palès que els diferents tipus de discriminacions poden relacionar-se entre si, i en el moment de fer-ho, generen altres situacions també discriminatòries (diferents que en el cas de no haver-se relacionat entre si).

2.2. Reconeixement de la discriminació interseccional

S'entén interseccionalitat com la manera per la qual diferents tipus de discriminació es relacionen o s'afecten entre si⁸¹³, i s'entén com a discriminació múltiple (també anomenada discriminació interseccional) qualsevol combinació de formes de discriminació contra persones per motiu de sexe, origen ètnic o racial, religió o creença,

⁸¹¹ GERARDS, J. i XENIDIS, R., *Algorithmic Discrimination in Europe...*, *op. cit.*, 2021, pàg. 116.

⁸¹² AÑÓN ROIG, M.J., "Desigualdades algorítmicas...", *op. cit.*, 2022, pàg. 22.

⁸¹³ Intersectionality. Cambridge Dictionary. Cambridge University Press. 2023. Disponible a: <https://dictionary.cambridge.org/es/diccionario/ingles/intersectionality>. Darrera visita el 2 de setembre de 2023.

discapacitat, edat, orientació sexual, identitat de gènere o altres característiques, i també aquella patida per persones que tenen o es percep que tenen aquestes característiques⁸¹⁴. Per tant, la interseccionalitat és una eina per a l'anàlisi de la discriminació⁸¹⁵, i la discriminació múltiple fa referència a la persona que es veu discriminada per més d'un motiu protegit. Com expressa Rey Martínez, no es tracta només que es creuin dos factors (per exemple, el sexe i l'ètnia) en una situació concreta, sinó el fet realment crucial que aquesta combinació de factors genera un conjunt nou, diferent dels anteriors⁸¹⁶. Que no s'adopti una perspectiva basada en la interseccionalitat per abordar la discriminació fa més difícil la detecció i reconeixement de discriminació múltiple. A més, la complexitat i la multidimensionalitat de les desigualtats (una realitat que es reflecteix en les dades que entrenen els algoritmes), col·lideix amb l'organització de la regulació antidiscriminació en base a atributs protegits únics i delimitats⁸¹⁷. Malgrat que des de la UE s'hagin intentat fer avenços en el reconeixement de la discriminació interseccional⁸¹⁸, la regulació antidiscriminació europea no estableix una obligació clara per redreçar-la⁸¹⁹ i el TJUE no la reconeix.

D'una banda, un exemple en el que no s'empra la perspectiva interseccional és en el cas Parris⁸²⁰, on el TJUE estableix que la disposició impugnada no és capaç de crear

⁸¹⁴Multiple discrimination. European Institute for Gender Equality. 2023. Disponible a: <https://eige.europa.eu/thesaurus/terms/1297>. Darrera visita el 2 de setembre de 2023.

⁸¹⁵ AÑÓN ROIG, M. J., “Principio Antidiscriminatorio...”, *op. cit.*, 2013, pàg. 151.

⁸¹⁶ REY MARTÍNEZ, F., “La discriminación múltiple, una realidad antigua, un concepto nuevo”, *Revista española de derecho constitucional*, núm. 84, 2008, pàg. 266 i 267; REY MARTÍNEZ, F., “O organismo español para a igualdad: claves para a súa converxencia co modelo europeo”, *Revista Galega de Administración Pública*, núm 61, 2021, pàg. 314-315. L'autor referencia com el TC no ha reconegut encara la discriminació múltiple. Posa d'exemple les sentències 69/2007, de 16 d'abril de 2007 i 1/2021, de 25 de gener de 2021, on el tribunal no va apreciar l'existència de discriminació. Defensa que, en canvi, una lectura en clau interseccional dels casos hagués evidenciat més fàcilment l'existència de discriminació per la combinació de l'ètnia i el gènere de la persona demandant.

⁸¹⁷ XENIDIS, R., “Turning EU equality law to algorithmic discrimination...”, *op. cit.*, 2020, pàg. 739.

⁸¹⁸ Vegeu: PARLAMENTO EUROPEO, *Resolución del Parlamento Europeo, de 21 de enero de 2021, sobre la estrategia de la Unión para la igualdad de género*, *op. cit.*, 2021, considerando D; COMISIÓN EUROPEA, *Una Unión de la igualdad: Estrategia para la Igualdad de Género 2020-2025*, COM(2020) 152 final, Bruselas, 5 de marzo de 2020, pàg. 2.

⁸¹⁹ XENIDIS, R., “Turning EU equality law to algorithmic discrimination...”, *op. cit.*, 2020, pàg. 741. Només es fa referència a la discriminació múltiple en el considerant 14 de la Directiva 2000/43/EC i el considerant 3 de la Directiva 2000/78/EC. En la Directiva 2004/113/EC i la Directiva 2006/54/EC no s'hi fa cap menció. Per tant, actualment no es dona resposta a aquesta problemàtica i les situacions de discriminació (combinada) no es protegeixen.

⁸²⁰TJUE (Sala Primera), sentència de 24 de novembre de 2016, C-443/15, David L. Parris contra Trinity College Dublin, Higher Education Authority, Department of Public Expenditure and Reform, Department of Education and Skills. La norma objecte d'estudi permet a les parelles unides civilment a rebre la pensió de viudetat si la unió s'havia format abans l'edat de 60 anys. El Sr. Parris no va poder contraure aquesta unió abans els 60 anys –malgrat estar en al relació durant 30 anys–, ja que les unions civils entre persones del mateix gènere no estaven permeses encara. Vegeu també, en el mateix sentit: TJUE (Gran Sala),

discriminació com a resultat de l'efecte combinat de l'orientació sexual i l'edat, ja que la norma no és discriminatòria aïlladament ni en base a l'orientació sexual ni en base a l'edat⁸²¹. De manera que s'analitzen les dues raons separadament per valorar si el Sr. Parris ha sofert un desavantatge particular. Aquesta sentència s'ha vist criticada en base al seu efecte perjudicial envers persones afectades per discriminació múltiple. D'altra banda, un cas en el qual s'empra aquesta perspectiva el sentència el TEDH, en l'assumpte Carvalho Pinto, on conclou que la qüestió que es planteja no és en relació amb l'edat o el sexe com a tals, sinó en la suposició que la sexualitat no és tan important per a una dona de cinquanta anys i mare de dos fills que per a una persona menor d'edat; evidenciant la suposició que la sexualitat femenina està essencialment vinculada a les finalitats de procreació i sense tenir en compte la seva importància física i psicològica per la realització personal de les dones⁸²². Es pot comprovar en aquesta reflexió, com el TEDH aposta per analitzar la situació des de la perspectiva de les dues categories sensibles que es veuen implicades en aquest cas, com són l'edat i el gènere, i com el fet de tenir-les en compte juntes suposa que s'abordi el cas apreciant-ne millor la seva complexitat⁸²³. També, per exemple, en un assumpte ocorregut a Finlàndia, el Tribunal Nacional de No Discriminació i Igualtat va investigar per quina raó a una persona se li havia retirat el crèdit per part d'un sistema de puntuació de crèdit algorítmic. El Tribunal va concloure que havia estat a causa que el sistema algorítmic va basar la seva decisió en múltiples caràcters com el gènere, l'edat i el lloc de residència (és a dir, la discriminació va ser el resultat de la combinació de diferents característiques), i això suposava un acte de discriminació il·legal⁸²⁴.

sentència de 18 de març de 2014, C-363/12, Z. contra A Government department and The Board of management of a community school.

⁸²¹En l'article 5.2 de la Llei 19/2020 s'especifica, en el mateix sentit que la sentència, que "*en cas de discriminació múltiple, la justificació de la diferència de tracte s'ha d'acomplir respecte de cadascun dels motius de discriminació*". En relació a altres casos en els quals el TJUE no ha emprat la perspectiva interseccional vegeu, per exemple: TJUE (Gran Sala), sentència de 14 de març de 2017, C-157/15, Samira Achbita i Centrum voor gelijkheid van kansen en voor racismebestrijding contra G4S Secure Solutions NV; TJUE (Gran Sala), sentència de 14 de març de 2017, C-188/15, Asma Bougnaoui i Association de défense des droits de l'homme (ADDH) contra Micropole SA.

⁸²²TEDH (Secció Quarta), sentència de 25 de juliol de 2017, 17484/15, Carvalho Pinto de Sousa Morais contra Portugal, par. 52.

⁸²³Vegeu també, en el mateix sentit: TEDH (Secció Tercera), sentència de 24 de juliol de 2012, 47159/08, B. S. contra Espanya, par. 71. El TEDH estima que les decisions dictades pels òrgans jurisdiccionals interns no van tenir en compte la vulnerabilitat específica de la demandant, inherent a la seva condició de dona africana exercint la prostitució, i aprecia una violació de l'article 14 del CEDH.

⁸²⁴Tribunal Nacional Finlandès de No Discriminació i Igualtat. Resum de la decisió disponible a: <https://www.yvtltk.fi/en/index/opinionsanddecisions/decisions.html#> Darrera visita el 2 de setembre de 2023.

La rellevància de reconèixer la discriminació múltiple quan ens referim a discriminació algorítmica té la seva base en els indicadors i les correlacions estadístiques. Com ja hem esmentat en l'apartat 1.2 del capítol II, un algoritme rarament discriminarà de manera oberta i directa per raó d'un atribut protegit, ja que usualment basarà els seus resultats en una multitud de factors, les bases de dades seran diverses i serà difícil relacionar la discriminació amb un atribut protegit concret, és a dir, serà difícil identificar exactament quines variables expliquen un resultat algorítmic concret i més complicat encara detectar i aïllar el motiu de discriminació⁸²⁵. La quantitat de dades creuades suposa un augment de la personalització dels resultats, i en conseqüència, un augment de les respostes específiques donades a aquelles persones en les quals concorren diverses categories que les fan especialment vulnerables⁸²⁶.

Per això, és probable que la noció de discriminació algorítmica interseccional tingui un major potencial per a generar avenços en matèria de prova si aquesta pot practicar-se tenint en compte els vincles de correlació⁸²⁷. Reconèixer la discriminació interseccional com una categoria doctrinal alliberària la càrrega de la prova en litigis de discriminació en fer possible que en casos en els quals cap tipus de discriminació pugui ser evidenciada per motius protegits separatament, es pugui evidenciar en base a diverses categories en combinació⁸²⁸, és a dir, en comptes d'aportar indicis de cada un dels atributs concurrents, es podria aportar un únic indicatiu interseccional⁸²⁹, donant l'oportunitat al tribunal d'analitzar els efectes de la combinació de categories en subgrups específics desavantatjats. D'aquesta manera, si determinades decisions no comporten un tractament menys favorable en relació a totes les dones respecte de tots els homes o a totes les persones d'una raça o grup ètnic minoritari respecte de la raça o grup ètnic majoritari, podria comportar tractament menys favorable si es troben en una situació comparable

⁸²⁵GILLIS, T.B. i SPIESS, J.L., “Big Data and Discrimination”, *University of Chicago Law Review*, vol. 86, 2019, pàg. 478; GERARDS, J. i XENIDIS, R., *Algorithmic Discrimination in Europe...*, *op. cit.*, 2021, pàg. 65; RIVAS VALLEJO, P., “Sesgos de género en el uso...”, *op. cit.*, 2022, pàg. 80. L'autora precisa que en la mesura en què l'aprenentatge automàtic no permet conèixer la correlació entre dades i quines d'elles han determinat el resultat, no és possible saber quins atributs analitzats han estat determinants pel resultat ofert pel sistema algorítmic, i si tècnicament no es factible donar aquesta explicació a la víctima de la decisió discriminatòria, la conseqüència és que probablement estiguem davant d'un cas de discriminació múltiple, interseccional, però no sigui possible constatar-ho llevat que la persona que contracta proporcioni dades de contrast.

⁸²⁶SORIANO ARNANZ, A., “Decisiones automatizadas y discriminación...”, *op. cit.*, 2021, pàg. 32.

⁸²⁷AÑÓN ROIG, M.J., “Desigualdades algorítmicas...”, *op. cit.*, 2022, pàg. 43 i 44.

⁸²⁸XENIDIS, R., “Turning EU equality law to algorithmic discrimination...”, *op. cit.*, 2020, pàg. 745.

⁸²⁹RIVAS VALLEJO, P., “Sesgos de género en el uso...”, *op. cit.*, 2022, pàg. 61.

totes les dones racialitzades i tots els homes blancs⁸³⁰. També, en relació amb la valoració de prova des de l'enfocament interseccional, el tribunal hauria d'anar més enllà, examinant quina tipologia de prejudicis causa la norma i sobre quins subjectes, i estudiant els efectes socials (d'ordre psicològic, econòmic, cultural) que poguessin afectar l'estatus material de les persones o el seu estatus social⁸³¹. Nogensmenys, tot i que es reconegué la discriminació múltiple en texts legislatius, hi ha limitacions a l'ús de la llei per promoure la igualtat, i més una igualtat sensible a la discriminació interseccional. La lògica d'igualtat desenvolupada mitjançant mecanismes judicials es limita a un model individual⁸³², mal equipat per implementar el concepte grupal d'igualtat i per abordar aspectes estructurals més complexos de discriminació, ocasionant que es miri a polítiques del sector públic com a solució per abordar aquestes tipologies de discriminació⁸³³.

L'anàlisi massiva de dades seguirà exacerbant les desigualtats existents fins que no s'estableixin sistemes per avaluar quin grau de desigualtat és tolerable⁸³⁴. La recerca de patrons, tendències o regles que expliquin el comportament de dades en un determinat context, objectiu de l'anàlisi massiva de dades, bàsicament reproduïx sistemes de subordinació, opressió i exclusió. Des de la perspectiva antidisubordinatòria, sempre que les dades estiguin lligades a comportaments socials, els resultats que s'obtinguin a partir d'elles seran discriminatoris, atès que simplement reflectiran una societat desigual i discriminatòria i opressora de grups de població determinats que comparteixen certs atributs protegits. Partint d'aquesta base, la normativa que s'instauri per regular la discriminació algorítmica ha de tenir en compte que si els sistemes algorítmics han estat entrenats amb dades que reflecteixen la societat, tant actual com passada, i per persones que viuen en la societat actual, mai podran no ser discriminatoris.

En aquest sentit, una mesura possible és la prohibició de l'ús de sistemes d'IA en contextos en els quals no es pugui assegurar que no s'estan perpetuant sistemes d'opressió i subordinació. També, la regulació podria, de manera més pràctica, assegurar que si més no, la discriminació que produeixen els sistemes algorítmics no és més exacerbada que aquella que produeixen les persones. Per exemple, com a control *ex ante*, establir tants

⁸³⁰ SORIANO ARNANZ, A., "Decisiones automatizadas y discriminación...", *op. cit.*, 2021, pàg. 31 i 32.

⁸³¹ AÑÓN ROIG, M. J., "Principio Antidiscriminatorio...", *op. cit.*, 2013, pàg. 139 i 140.

⁸³² SQUIRE, J., "Intersecting inequalities", *International Feminist Journal of politics*, vol. 11, núm. 4, 2009, pàg. 507.

⁸³³ MABBETT, D., "Aspirational Legalism and the Role of the Equality and Human Rights Commission in Equality Policy", *The Political Quarterly*, vol 79, núm. 1, 2008, pàg. 46.

⁸³⁴ BAROCAS, S. i SELBST, A. D., "Big Data's Disparate...", *op. cit.*, 2016, pàg. 728.

per cent de discriminació tolerable -és a dir, si l'algoritme no és un 95% precís per tots els grups de població -tenint en compte també la interseccionalitat (és a dir, per exemple, tenint en compte un grup de població que siguin dones, afroamericanes de més de 60 anys)- no es podrà introduir en el mercat. En el mateix sentit, però establint un altre sistema de cribratge, la normativa podria establir que no es poden emprar sistemes en els quals les taxes d'error per a tots els grups (incloent categories protegides combinades) superés l'1%. Per exemple, Buolamwini i Gebru presenten un mètode per avaluar el biaix present en conjunts de dades i algoritmes d'anàlisi facial automatitzats respecte als subgrups fenotípics. S'avaluen tres sistemes comercials de classificació en base al gènere, que mostren que les dones negres són el grup més mal classificat (amb taxes d'error fins al 34,7%), quan, contràriament, la màxima taxa d'error per homes blancs és del 0,8%⁸³⁵. Llavors, en existir disparitats substancials en l'exactitud del sistema segons els trets físics de les persones sobre les quals s'aplica, aquests no podrien emprar-se.

A tall de recapitulació, per pal·liar el fet que la normativa antidiscriminació no atorgui suficient rellevància a l'element grupal definitori no eliminable de la discriminació, és necessària la inclusió de la interseccionalitat i la discriminació estructural en l'anàlisi dels casos. Quant a la perspectiva en base a la discriminació estructural, pot ajudar a interpretar el tracte que produeix una discriminació a la llum de la posició social estructural de subordinació, no obstant que es manifesti de forma individualitzada; a entendre que els algoritmes repliquen i perpetuen les estructures de dominació i que la perspectiva de la discriminació com a greuge individual que presenta majoritàriament la normativa antidiscriminació no és suficient; i, anant més enllà, a instrumentalitzar el dret com a eina transformadora de les estructures socials creadores dels sistemes d'opressió amb una visió centrada en la lògica subjacent a l'ús de la intel·ligència artificial. Quant a la perspectiva interseccional, pot ajudar a millorar la interpretació del resultat d'una decisió presa mitjançant un sistema algorítmic, que en molts casos, no basarà la discriminació en un atribut protegit, sinó que ho farà en base a múltiples atributs protegits. Aquesta perspectiva pot possibilitar que es tingui en compte que si determinades decisions automatitzades no comporten un tractament menys favorable en relació amb un atribut protegit, sí que ho poden fer en base a diversos atributs protegits combinats.

⁸³⁵BUOLAMWINI, J., i GEBRU, T., "Gender Shades: Intersectional Accuracy...", *op. cit.*, 2018, pàg. 77.

3. Flexibilització dels conceptes discriminació directa i discriminació indirecta

El TJUE a través de la seva jurisprudència ha anat perfilant els contorns dels conceptes de discriminació directa i indirecta, i inclús ha introduït el concepte de discriminació per associació. Tal com en l'apartat tercer i quart del capítol II hem analitzat les problemàtiques que presentava la discriminació algorítmica per ser enquadrada tant en la conceptualització de la discriminació directa com de la discriminació indirecta, en aquest apartat presentem com a través de la jurisprudència del TJUE hi ha certs elements que poden emprar-se en els casos de discriminació algorítmica. En primer lloc plantejem els elements d'interès de la jurisprudència de cada tipologia de discriminació (3.1), i en segon lloc abordem la jurisprudència relativa a la discriminació per associació i el seu interès per regular la discriminació algorítmica per associació i per error (3.2).

3.1. Elements d'interès de cada tipologia de discriminació

En ambdós casos és favorable per abordar la discriminació algorítmica el fet que no sigui rellevant la intencionalitat per poder portar un cas de discriminació a litigi⁸³⁶. En el cas dels algoritmes, per a la persona que ha patit discriminació és més fàcil poder obviar si aquesta ha estat o no intencional, essent suficient demostrar que una categoria protegida juga un paper en el tractament algorítmic que resulta en el desavantatge d'un grup protegit.

A part d'aquest element que comparteixen les dues tipologies de discriminació, hi ha elements d'interès per la discriminació algorítmica que només es donen en casos de discriminació directa (3.1.1) i, d'igual manera, hi ha elements propis de casos de discriminació indirecta (3.1.2).

⁸³⁶ GERARDS, J. i XENIDIS, R., *Algorithmic Discrimination in Europe...*, op. cit., 2021, pàg. 67 i 70.

3.1.1. Elements d'interès de la discriminació directa

En aquest apartat analitzarem quins elements de la discriminació directa poden ser útils quan ens referim a la discriminació algorítmica. Aquests elements es concreten en la possibilitat d'obviar una referència comparativa com s'estableix en el cas Dekker (3.1.1.1), i la consideració de discriminació directa no en base a l'atribut protegit, sinó en base a elements indissociables a l'atribut protegit, com s'estableix en el cas Tadao Maruko (3.1.1.2).

3.1.1.1. Obviar una referència comparativa en la discriminació directa: Cas Dekker

En el cas Dekker, l'any 1981, una candidata a una feina d'educadora va ser rebutjada atès que va comunicar durant el procés selectiu que estava embarassada⁸³⁷. El tribunal nacional va preguntar al TJUE, si era lícita una negativa de contractació per raó d'embaràs en relació amb el principi d'igualtat de tracte en matèria d'accés al treball.

L'advocat general va exposar que la negativa de contractar una candidata deguda la seva imminent maternitat, en tenir en compte un fet que afectava exclusivament a les dones treballadores, constituïa una discriminació directa per raó de sexe⁸³⁸. A partir d'aquest cas es va iniciar una modificació relativa a l'exigència de comparació en els litigis de discriminació per raó d'embaràs. Actualment, una llarga jurisprudència del TJUE estableix clarament que si el perjudici patit per una persona és per raó del seu embaràs, la situació es qualifica de discriminació directa, sense necessitat de referència comparativa⁸³⁹.

⁸³⁷TJCE, sentència de 8 de novembre de 1990, C-177/88, Elisabeth Johanna Pacifica Dekker contra Stichting Vormingscentrum voor Jong Volwassenen (VJV-Centrum) Plus, par. 2.

⁸³⁸ Conclusiones del abogado general Marco Darmon, presentadas el 14 de noviembre de 1989, en el asunto C-177/88, Elisabeth Johanna Pacifica Dekker contra Stichting Vormingscentrum voor Jong Volwassenen (VJV-Centrum) Plus, par. 25.

⁸³⁹ ANÓN ROIG, M. J., "Principio Antidiscriminatorio...", *op. cit.*, 2013, pàg. 143 i 144.

La classificació com a discriminació directa va venir determinada perquè la negativa de contractació per raó d'embaràs només es podia aplicar a candidates de sexe femení⁸⁴⁰. No hi hauria situació comparable atès que els candidats de sexe masculí mai podrien rebre una negativa de contractació basada en l'embaràs. Per tant, s'establia l'embaràs com factor indissociable al sexe femení.

Aquesta línia jurisprudencial sembla favorable per abordar la discriminació algorítmica arran de factors equivalents a l'atribut protegit. Tanmateix, cal apuntar que es pot aplicar sota la condició que hi hagi una transparència mínima en el sistema algorítmic que permeti determinar en base a quins factors s'ha pres la decisió, per poder establir si aquests són o no són indissociables a l'atribut protegit.

A continuació analitzarem el cas Tadao Maruko, que també tracta sobre els elements indissociables a l'atribut protegit.

3.1.1.2. Generar nexa causal en la discriminació directa per elements inseparables del motiu protegit: Cas Tadao Maruko

En el cas Tadao Maruko, el TJUE havia de respondre si una disposició estatutària d'un organisme alemany amb personalitat jurídica pública que gestionava assegurances era contrària a l'article 2.2.a de la Directiva 2000/78/CE, que disposava que un membre d'una parella inscrita, en morir l'altre membre de la parella, no tenia dret a percebre una pensió de supervivència tal i com correspondria a un cònjuge, malgrat haver mantingut una unió similar al matrimoni.

L'advocat general va afirmar que denegar -en base a la disposició estatutària- la pensió, perquè no hi havia matrimoni (reservat a persones heterosexuales), quan s'havia formalitzat una unió, amb efectes substancialment idèntics, entre persones del mateix sexe, suposava una discriminació indirecta per raó d'orientació sexual⁸⁴¹. Segons l'advocat general no es tractava d'una discriminació directa, ja que el rebuig no es

⁸⁴⁰ TJCE, sentència de 8 de novembre de 1990, C-177/88, Elisabeth Johanna Pacifica Dekker contra Stichting Vormingscentrum voor Jong Volwassenen (VJV-Centrum) Plus, par. 10.

⁸⁴¹ Conclusiones del abogado general Dámaso Ruiz-Jarabo Colomer, presentadas el 6 de septiembre de 2007, en el asunto C-267/06, Tadao Maruko contra Versorgungsanstalt der deutschen Bühnen, par. 102.

fundava en l'orientació sexual de la persona interessada, sinó que es fundava en una disposició aparentment neutra -distinció entre parelles inscrites i parelles casades-, que ocasionava un desavantatge a persones homosexuals⁸⁴².

Nogensmenys, el TJUE va obrir la possibilitat a que el tribunal nacional valorés si es tractava de discriminació directa en establir que si només es podia ser parella inscrita quan es tractava de la unió de persones del mateix sexe, i la disposició negava la pensió de supervivència a causa de ser parella inscrita, s'estava discriminant directament a causa de l'orientació sexual⁸⁴³. Alemanya va crear la possibilitat d'establir-se com a parella inscrita, per adaptar el seu ordenament jurídic i permetre la unió de persones del mateix sexe, havent optat per no obrir la institució del matrimoni, mantenint-se reservada per persones de diferent sexe. Per tant, es va crear un règim diferent, que progressivament va anar assimilant les condicions aplicables al matrimoni⁸⁴⁴. Llavors, la parella inscrita, sense ser idèntica al matrimoni, gaudia d'una situació comparable en allò relatiu a la prestació de supervivència, que la disposició estatutària controvertida negava a les parelles inscrites⁸⁴⁵.

En aquest cas es va més enllà que en el cas Dekker, ja que el nexa causal entre motiu protegit i disposició, pràctica o criteri no era tan estreta. No es tractava d'una situació en la qual només es podrien trobar les dones (embaràs), sinó que era una disposició la que creava la disparitat. Tots els beneficis que s'associessin al matrimoni no podien ser gaudits per aquelles persones que no poguessin accedir a la institució del matrimoni, de manera que el matrimoni estava indissociablement lligat a l'orientació sexual, entrant en un cas, llavors, de discriminació directa.

Aquesta interpretació extensiva de l'existència d'un nexa causal seria interessant per abordar les situacions en què l'indicador i l'atribut protegit pràcticament se superposen. Ha d'existir un nexa causal entre el tracte menys favorable i l'atribut protegit, tot i que la disposició o pràctica no ha de fer necessàriament referència explícita a l'atribut protegit, sempre que es mencioni un altre factor indissociable d'aquest. La jurisprudència del TJUE

⁸⁴² *Ídem, ibidem*.

⁸⁴³ TJCE (Gran Sala), sentència d'1 d'abril de 2008, C-267/06, Tadao Maruko contra Versorgungsanstalt der deutschen Bühnen, par. 72.

⁸⁴⁴ *Ídem*, par. 67.

⁸⁴⁵ *Ídem*, par. 69-72.

considera una acció constitutiva de discriminació directa quan el criteri formal en relació al qual es pren una decisió és inseparable de la pertinença al grup desavantajat⁸⁴⁶.

S'ha argumentat que en la discriminació mitjançant sistemes algorítmics, sota aquesta jurisprudència, malgrat l'opacitat de l'algoritme, es podria determinar que existeix discriminació directa⁸⁴⁷. Per exemple, quan un sistema algorítmic sobre el qual es desconeix l'element o elements que determinen la decisió presa, rebutja totes les dones i només contracta a homes (malgrat que totes les persones candidates compleixen uns requisits mínims per ocupar els llocs de feina). Si en el resultat de l'algoritme s'evidencia un tracte menys favorable per a tots els membres d'un grup determinat i especialment protegit, es podrà concloure que la decisió s'ha pres basant-se en la categoria especialment protegida o en elements de judici indissociables d'aquella⁸⁴⁸, i en aquest cas es podrà considerar l'existència de discriminació directa si es pren l'algoritme en la seva totalitat com a pràctica discriminatòria. Aquests casos s'auguren, així i tot, molt improbables.

Resumidament, els casos aquí presentats semblen interessants en flexibilitzar el concepte de discriminació directa perquè abasti el màxim de situacions possibles. Tanmateix, com ja hem plantejat en el cas Dekker, si no se sap quin procés decisor ha seguit el sistema algorítmic serà molt complicat establir quina ha estat la causa del resultat discriminatori, i per tant, establir un cas de discriminació directa. A més, no està clar quan el vincle entre les correlacions i els indicadors estadístics i un atribut protegit serà considerat suficientment consistent per part del TJUE per considerar el cas com discriminació directa⁸⁴⁹. Tal com en el cas Dekker⁸⁵⁰ el TJUE va tractar la discriminació per embaràs com a discriminació directa en establir un vincle indissociable entre el sexe femení i l'embaràs, en el cas Jyske Finans⁸⁵¹, va determinar que el país de naixement no servia per si sol per establir una presumpció general de pertinença a un determinat grup ètnic, és a dir, no va considerar l'existència d'un vincle directe o indissociable entre el país de naixement i l'origen ètnic. En aquest assumpte, el TJUE va demanar més d'un indicador

⁸⁴⁶ SORIANO ARNANZ, A., "Decisiones automatizadas y discriminación...", *op. cit.*, 2021, pàg. 16.

⁸⁴⁷ SORIANO ARNANZ, A., "Discriminación algorítmica: garantías...", *op. cit.*, 2022, pàg. 152 i 153.

⁸⁴⁸ *Ídem, ibidem*.

⁸⁴⁹ XENIDIS, R., "Turning EU equality law to algorithmic discrimination...", *op. cit.*, 2020, pàg. 746.

⁸⁵⁰ TJUE, sentència de 8 de novembre de 1990, C-177/88, Elisabeth Johanna Pacifica Dekker contra Stichting Vormingscentrum voor Jong Volwassenen (VJV-Centrum) Plus, par. 2.

⁸⁵¹ TJUE (Sala Primera), sentència de 6 d'abril de 2017, C-668/15, Jyske Finans A/S contra Ligebehandlingsnavnet, par. 20.

del motiu protegit, considerant que un únic indicador no era suficient per substituir-lo⁸⁵² (aquí hagués estat summament rellevant, com planteja Lilla Farkas, una anàlisi interseccional)⁸⁵³.

3.1.2. Elements d'interès de la discriminació indirecta

En aquest apartat analitzem quins elements de la discriminació indirecta poden ser útils quan ens referim a la discriminació algorítmica. Estudiem com el fet que la discriminació indirecta s'estableixi en base als efectes de la disposició, pràctica o criteri discriminatoris -com queda patent en els casos Jenkins i Bilka-Kaufhaus que inicien una jurisprudència del TJUE relativa a la discriminació retributiva per raó de sexe- pot servir per captar la discriminació algorítmica (3.1.2.1), i com altres factors poden contribuir a que sigui un refugi pels casos en què, segurament, no es podrà establir que ha existit discriminació directa (3.1.2.1).

3.1.2.1. Enfocament en els resultats: Cas Jenkins i Bilka- Kaufhaus

En el cas Jenkins, una empresa va fixar la retribució per hora pel treball a temps parcial en una quantia inferior en un 10% a la retribució de treball a temps complet. La demandant al·legava que aquesta decisió afectava més desfavorablement a les dones, ja que treballaven més a temps parcial⁸⁵⁴. El TJUE va determinar que la diferència de retribució en l'hora a temps complet i a temps parcial només constituïa discriminació per raó de sexe en el cas que en realitat fos un mitjà indirecte per reduir la retribució de persones

⁸⁵² *Ídem*, par. 17 i 18. Defensa el TJUE que el país de naixença només seria un dels factors específics que permetrien concloure que una persona pertany a un grup ètnic, sense ser l'únic.

⁸⁵³ FARKAS, L., *The Meaning of Racial or Ethnic Origin in EU Law: Between Stereotypes and Identities*, European Network of Legal Experts in Gender Equality and Non-Discrimination, Brussels, 2017, pàg. 112.

⁸⁵⁴ TJCE, sentència de 31 de març de 1981, C- 96/80, J.P. Jenkins contra Kingsgate (Clothing Productions) Ltd, par. 2 i 6.

treballadores a temps parcial, atès que el grup estava integrat de manera exclusiva o preponderant per dones⁸⁵⁵.

En el cas Bilka-Kaufhaus, una treballadora va impugnar la denegació de la pensió d'empresa al·legant que el règim de pensions infringia el principi d'igualtat entre treballadors i treballadores perquè l'empresa exigia un mínim de treball a jornada completa, fet que perjudicava les dones, que donades les càrregues familiars, es veien obligades amb major freqüència que els treballadors a optar pel treball a temps parcial⁸⁵⁶.

El TJUE va determinar que quan resultés que un percentatge considerablement inferior de dones que d'homes treballava a jornada completa, l'exclusió de persones treballadores a temps parcial del règim de pensions de l'empresa era discriminació per raó de sexe, atès que, tenint en compte les dificultats trobades per les treballadores per treballar a jornada completa, aquesta mesura no podia explicar-se amb factors que excloguessin la discriminació per raó de sexe⁸⁵⁷. L'advocat general raonava, que les pressions socioculturals que pesen sobre les dones treballadores no havien de ser explotades per l'empresariat⁸⁵⁸.

Actualment, tal com passava en els casos esmentats, els elements que determinen la diferència de remuneració, per exemple, en el cas de plataformes col·laboratives, no són inherents al sexe però estan relacionats amb la manera amb què es construeixen els rols de gènere, és a dir, el fet que les dones segueixin suportant una major part de la responsabilitat de cura de familiars que els homes continua afectant fortament les seves carreres i la seva posició econòmica⁸⁵⁹. Soriano Arnanz considera que es podria argumentar que la manera com es calculen, a través de sistemes algorítmics, els salaris en les plataformes de transport col·laboratiu, perjudica de manera desproporcionada les dones, aplicant-se la jurisprudència a partir de l'assumpte Bilka-Kaufhaus⁸⁶⁰.

⁸⁵⁵ *Ídem*, par. 15.

⁸⁵⁶ TJCE, sentència de 13 de maig de 1986, C-170/84, Bilka - Kaufhaus GmbH contra Karin Weber von Hartz, par. 5-6.

⁸⁵⁷ *Ídem*, par. 29.

⁸⁵⁸ Conclusiones del abogado general Marco Darmon, presentadas el 15 de octubre de 1985, en el asunto C-170/84, Bilka - Kaufhaus GmbH contra Karin Weber von Hartz, par. 13.

⁸⁵⁹ ARABADJIEVA, K. i ZWYSEN, W., *Gender inequality in performance-related pay: a gap in the EU equal pay agenda*, ETUI Research Paper - Policy Brief, Brussels, February 2022, pàg.7.

⁸⁶⁰ SORIANO ARNANZ, A., "Discriminación algorítmica: garantías...", *op. cit.*, 2022, pàg. 162.

No obstant això, sembla difícil aïllar concretament la causa de la diferència de salaris. Pot succeir, per exemple, que es tracti d'un sistema retributiu caracteritzat per una falta absoluta de transparència, de manera que només es pugui arribar a la conclusió que els homes cobren més que les dones. Llavors, el TJUE s'inclina per atribuir la càrrega de la prova a la persona empresària, que ha d'acreditar el caràcter neutre del sistema retributiu aplicat⁸⁶¹. En aquest cas, llavors, les treballadores de plataformes col·laboratives, si tinguessin informació grupal que acredités que cobren menys que els homes que treballen en la mateixa plataforma, podrien litigar per estar patint discriminació indirecta. Tanmateix, com dèiem en l'apartat segon del capítol II, en aquests casos no es tracta de fets aïllats sinó de discriminacions estructurals. Les dones víctimes de discriminació retributiva no litiguen perquè estan aïllades i no tenen accés a les dades que sustenten les seves possibles demandes. Les discriminacions retributives no constitueixen fets aïllats i desviats de la normalitat, sinó que són pautes de comportament molt arrelades socialment⁸⁶².

A continuació abordem com la discriminació indirecta també pot útil per captar la discriminació algorítmica si no es pot establir la relació entre indicador i atribut protegit.

3.1.2.2. Refugi en la discriminació indirecta si no es pot establir la relació entre indicador i atribut protegit

Tal com hem plantejat, la discriminació indirecta en comptes de focalitzar-se en el tractament de diverses persones basat en la pertinença a un grup, es concentra en els efectes de qualsevol decisió, mesura o política desavantajosos per a un grup protegit de persones. Això genera que, en enfocar-se en els efectes, pot evadir els obstacles en relació al contingut de l'algoritme, la comprensió per part de persones que no són expertes de com funciona el sistema i les qüestions de responsabilitat, des de persones

⁸⁶¹ CABEZA PEREIRO, J., "La discriminación retributiva por razón de sexo como paradigma de la discriminación sistémica", *Lan Harremanak-Revista de relaciones laborales*, núm. 25, 2011, pàg. 91. Vegeu, per exemple, el cas ja esmentat anteriorment: TJUE, sentència de 17 d'octubre de 1989, C-109/88, Handels-og Kontorfunktionærernes Forbund i Danmark y Dansk Arbejdsgiverforening, en nom de Danfoss, par. 11.

⁸⁶² *Ídem*, pàg. 97 i 98.

desenvolupadores fins a usuàries finals⁸⁶³. A més, la discriminació indirecta, en focalitzar-se en els criteris, pràctiques o disposicions que -malgrat la seva aparença de neutralitat- són discriminatoris, és una doctrina que s'adapta a aquelles discriminacions que trobem en els algoritmes si les dades emprades per programar-lo o entrenar-lo reflecteixen biaixos i estereotips que s'han cristal·litzat en patrons de desigualtat⁸⁶⁴.

El concepte tal com és interpretat pel TJUE, és capaç d'adreçar millor les situacions de discriminació arran d'indicadors. Per consegüent, sembla millor posicionat per abordar la discriminació en base a indicadors estadístics quan es debat si la vinculació entre l'indicador i l'atribut protegit no és suficientment evident com per considerar-se discriminació directa. Seria el cas de situacions en què un algoritme emprava les dades de residència i codi postal per inferir l'origen ètnic de les persones i discriminar-les per aquesta raó⁸⁶⁵.

Tanmateix, Sáez Lara qüestiona la fàcil equiparació de la discriminació algorítmica amb un supòsit de discriminació indirecta (en l'àmbit laboral però equiparable a altres àmbits), atès que quan s'ha de fer front a una decisió automatitzada el problema desborda o supera el supòsit de fet donada la seva complexitat, i suggereix una nova ampliació de la tutela antidiscriminatòria⁸⁶⁶. També ho planteja Añón Roig, afirmant que no obstant que la discriminació indirecta es proposa com a paraigües jurídic per afrontar la discriminació algorítmica, les dificultats que planteja aquesta última poden requerir la reconducció de la discriminació cap a modalitats diferents a la discriminació indirecta (específicament la discriminació per associació i la discriminació interseccional), atès que es considera que aquestes categories poden superar algunes deficiències del concepte de discriminació indirecta⁸⁶⁷. Com ja hem vist en el segon apartat d'aquest capítol, és cert que la discriminació interseccional pot aportar elements per millorar la protecció envers la discriminació algorítmica. Ens queda, llavors, seguint la proposta d'Añón Roig, abordar la discriminació per associació i per error.

⁸⁶³GERARDS, J. i XENIDIS, R., *Algorithmic Discrimination in Europe...*, *op. cit.*, 2021, pàg. 71.

⁸⁶⁴*Ídem*, pàg. 72.

⁸⁶⁵*Ídem*, pàg. 71.

⁸⁶⁶SÁEZ LARA, C., "El algoritmo como protagonista de la relación laboral...", *op. cit.*, 2020, pàg. 49.

⁸⁶⁷ AÑÓN ROIG, M.J., "Desigualdades algorítmicas...", *op. cit.*, 2022, pàg. 43 i 44.

3.2. Discriminació per associació i per error

Com hem explicat en l'apartat 2.2.2.5 del capítol I, la discriminació per associació és probable en casos de discriminació algorítmica. La discriminació per associació és aquella que es produeix quan una persona (o grup de persones), degut a la seva relació amb una altra persona (o grup) que ostenta un atribut protegit, pateix un tracte discriminatori⁸⁶⁸. La discriminació per associació es reconeix (no en contextos de sistemes algorítmics sinó en casos de discriminació tradicional) per part del TJUE en els assumptes Coleman⁸⁶⁹, CHEZ⁸⁷⁰ i Hakelbracht i altres⁸⁷¹. Exposem, en primer lloc, els casos Coleman (3.2.1), CHEZ (3.2.2), i Hakelbracht i altres (3.2.3), per analitzar seguidament què poden aportar a la regulació de la discriminació algorítmica (3.2.4).

3.2.1. Cas Coleman

En l'assumpte Coleman el TJUE va determinar que existia discriminació directa per raó de discapacitat, en el cas de l'assetjament i altres formes de discriminació, patit per la cuidadora (mare) d'una persona amb discapacitat per part de qui l'havia contractada. Va concloure que la prohibició de discriminació directa no es circumscriu exclusivament a aquelles persones que tenien discapacitat, sinó que, quan una persona que contractava tractava a una persona treballadora (que no tenia ella mateixa discapacitat) de manera menys favorable a com tractava, havia tractat, o podria tractar a una altra persona treballadora en una situació anàloga, i la primera acreditava que el tracte menys favorable que rebia estava motivat per la discapacitat d'un fill seu, a qui prodigava la major part de

⁸⁶⁸ La Llei 15/2022 defineix la discriminació per associació en el seu article 6.2.a. El Texto Refundido de la Ley General de derechos de las personas con discapacidad y de su inclusión social la defineix en l'article 2.e. També la Llei 19/2020 defineix aquest concepte en l'article 4, apartat o.

⁸⁶⁹TJCE (Gran Sala), sentència de 17 de juliol de 2008, C-303/06, S. Coleman contra Attridge Law and Steve Law.

⁸⁷⁰TJUE (Gran Sala), sentència de 16 de juliol de 2015, C-83/14, CHEZ Razpredelenie Bulgaria AD contra Komisia za zashtita ot diskriminatsia.

⁸⁷¹ TJUE (Sala Tercera), sentència de 20 de juny de 2019, C-404/18, Jamina Hakelbracht y otros contra WTG Retail BVBA.

les cures que requeria, el tracte resultava contrari a la prohibició de discriminació directa⁸⁷².

El TEDH també va confirmar que l'article 14 del CEDH⁸⁷³ cobria els casos en què una persona rebia un tracte menys favorable sobre la base d'una característica protegida d'una altra persona. En l'assumpte Guberina contra Croàcia⁸⁷⁴, per exemple, va estimar que el tracte discriminatori rebut pel demandant degut a la discapacitat del seu fill era una forma de discriminació per motiu de discapacitat.

En termes generals, es pot concloure llavors que la discriminació directa també pot donar-se quan una persona que no pertany a un grup especialment protegit resulta directament perjudicada per una decisió, conducta o mesura deguda a la seva relació amb els membres d'un grup protegit⁸⁷⁵.

L'advocat general Poiares Maduro establia en les seves conclusions sobre l'assumpte que col·locar directament en el punt de mira una persona per tenir una característica particular no era l'única manera de discriminar-la, en existir maneres més subtils i menys evidents de fer-ho. Establia que una sòlida concepció de la igualtat requeria que la legislació que reprimia la discriminació s'apliqués així mateix a aquelles formes més subtils de discriminació, atès que també afectaven a aquelles persones incloses en categories sensibles⁸⁷⁶. També afirmava que era possible ser víctima de discriminació il·legal per motiu de discapacitat en virtut de la Directiva 2000/78/CE sense que un/a fos discapacitat/da, establint que el que era rellevant consistia en què la discapacitat de la qual es tractés -en aquest cas la discapacitat del fill de la Sra. Coleman-, hagués sigut emprada per tractar pitjor a una persona⁸⁷⁷. El principal valor de la sentència Coleman és l'atenció a l'element objectiu (les causes de discriminació) com a via per l'ampliació de l'element subjectiu (les persones protegides) i, amb això, el pas de la tutela *rationae personae* a la

⁸⁷² TJCE (Gran Sala), sentència de 17 de juliol de 2008, C-303/06, S. Coleman contra Attridge Law and Steve Law, par. 56.

⁸⁷³ Article 14 CEDH.

⁸⁷⁴ TEDH (Secció Segona), sentència de 22 de març de 2016, 23682/13, Guberina contra Croàcia. Vegeu també, com a cas de discriminació per associació: TEDH (Secció Segona), sentència de 28 de març de 2017, 25536/14, Skorjanec contra Croàcia. En aquest cas, una dona va ser agredida juntament amb la seva parella pel seu origen gitano, origen que no compartia la seva parella.

⁸⁷⁵ SORIANO ARNANZ, A., "Discriminación algorítmica: garantías...", *op. cit.*, 2022, pàg. 148.

⁸⁷⁶ Conclusiones del abogado general Miguel Poiares Maduro, presentadas el 31 de enero de 2008, en el asunto C-303/06, S. Coleman contra Attridge Law i Steve Law, par. 12.

⁸⁷⁷ *Ídem*, par. 23.

tutela *rationae materiae*⁸⁷⁸. Es va acudir, en aquesta sentència als objectius i efecte útil de la Directiva 2000/78/CE per fonamentar la seva interpretació extensiva⁸⁷⁹.

Maneiro Vázquez va més enllà i determina que en aquest cas es podria haver apreciat, a més de la discriminació per discapacitat, una discriminació per raó de sexe i gènere en la mesura que l'afectada ho és, precisament per al seva condició de mare (sexe) i cuidadora (gènere) de la persona discapacitada⁸⁸⁰, de manera que la discriminació per associació produïda no ha d'estar moguda per una única causa, sinó que pot tractar-se d'una discriminació múltiple.

3.2.2. Cas CHEZ

En el cas CHEZ (com ja hem explicat breument en l'apartat 4.3.1 del capítol II), una dona -que va manifestar que no pertanyia a l'ètnia gitana-, que tenia un establiment obert al públic en el barri Gizdova Mahala de Bulgària -majoritàriament poblat per persones d'ètnia gitana-, va denunciar que els comptadors elèctrics en el barri estaven col·locats a una altura de 6 o 7 metres de terra, quan en els altres barris estaven a 1,5 metres i que això impedia que pogués visualitzar les dades del comptador. El tribunal va apreciar que malgrat que la demandant no fos d'ètnia gitana estava patint els efectes d'una norma implementada per donar un tracte menys favorable a persones d'ètnia gitana. Per tant, establia que el concepte “discriminació per raó d'origen ètnic” era aplicable amb independència que la mesura col·lectiva afectés a persones amb un origen ètnic específic, o les que, sense tenir aquest origen, patien amb les primeres el tracte menys favorable o el desavantatge particular derivat d'aquesta mesura⁸⁸¹.

Establia l'advocada general d'aquest assumpte que ni l'article 21 de la CDFUE, ni la Directiva 2000/43 restringien l'aplicació del principi d'igualtat de tracte a persones que fossin discriminades a causa de la “seva” (en el sentit de la seva pròpia) raça o origen

⁸⁷⁸ MANEIRO VÁZQUEZ, Y., *La discriminación por asociación: desafíos sustantivos y procesales. Desde su creación por el Tribunal de Justicia a su recepción por la doctrina judicial española*, Pamplona, Aranzadi, 2022, pàg. 61.

⁸⁷⁹ *Ídem*, pàg. 65.

⁸⁸⁰ *Ídem*, pàg. 106.

⁸⁸¹ TJUE (Gran Sala), sentència de 16 de juliol de 2015, C-83/14, CHEZ Razpredelenie Bulgaria AD contra Komisia za zashtita ot diskriminatsia, par. 60.

ètnic. Les disposicions, redactades de forma més general, prohibien la discriminació per raó de raça o origen ètnic. Concloïa que aquesta redacció permetia invocar el principi de no-discriminació a persones que simplement fossin codiscriminades⁸⁸². Al dirigir-se, la pràctica controvertida, de forma general i col·lectiva, contra totes les persones que a Gizdova Mahala es veien subministrades d'electricitat per l'empresa, si es constata que la pràctica estava associada a una discriminació contra les persones d'ètnia gitana que vivien en el barri, el caràcter general i col·lectiu de la pràctica donava lloc a que inevitablement també fossin codiscriminades les persones que no eren d'ètnia gitana, atès que havien de patir l'efecte estigmatitzant de la mesura general i col·lectiva tant com les persones d'ètnia gitana⁸⁸³.

En l'assumpte CHEZ, malgrat que la demandant no era d'ètnia gitana, el TJUE va entendre que el principi de no-discriminació tenia la intenció de beneficiar a persones que, no obstant no ser membres d'un grup protegit, patien un tractament més desfavorable o un desavantatge particular en base a un atribut protegit⁸⁸⁴. La sentència CHEZ, va estendre l'àmbit subjectiu de la discriminació per associació respecte del que havia precisat anteriorment la sentència Coleman. En aquesta, la persona discriminada per associació es caracteritzava per guardar una relació personal estreta amb la persona discapacitada i tenir, a més, la condició de cuidadora. En canvi, la sentència CHEZ ja no exigeix que el subjecte discriminat per associació guardi “una estreta relació personal” amb qui presenta la característica discriminatòria, de manera que això suposa que la persona que es veu discriminada no té perquè compartir la característica, podent ser víctima de discriminació per associació quan comparteixi danys col·laterals com a conseqüència d'una mesura discriminatòria de caràcter general i col·lectiu, sense necessitat que s'acrediti una estreta vinculació amb terceres persones⁸⁸⁵.

⁸⁸² Conclusiones de la abogada general Juliane Kokott, presentadas el 12 de marzo de 2015, en el asunto C-83/14, CHEZ Razpredelenie Bulgaria AD contra Komisia za zashtita ot diskriminatsia, par. 53-54.

⁸⁸³ *Ídem*, par. 60.

⁸⁸⁴ XENIDIS, R., “Turning EU equality law to algorithmic discrimination...”, *op. cit.*, 2020, pàg. 748. L'autora exposa que aquesta interpretació s'ajusta al considerant 16 i a l'article 3.1 de la Directiva 2000/43/CE, en el qual s'explica que la discriminació per raó de raça o origen ètnic que ha de garantir la directiva ha de beneficiar a totes les persones.

⁸⁸⁵ MANEIRO VÁZQUEZ, Y., *La discriminación por asociación...*, *op. cit.*, 2022, pàg. 75-76.

En aquest cas també es va establir que, inclús si un interès legítim existia i els mitjans eren necessaris, la pràctica es podria considerar igualment discriminatòria atès que es valorava com “ofensiva i que estigmatitzava” i per tant, desproporcionada⁸⁸⁶.

A més, és un cas interessant també en relació amb la distinció entre discriminació directa i indirecta. El tribunal va establir que es tractava de discriminació directa en el cas que el criteri neutral (posar els cables a 6 o 7 metres de terra) s’imposava per una raó d’origen ètnic (que el barri estava poblat majoritàriament per persones d’ètnia gitana). En el cas que el criteri no s’hagués establert per raó de l’ètnia de les persones que habitaven el barri, es tractaria de discriminació indirecta⁸⁸⁷. Per tant, en aquest context, establir l’existència de discriminació directa o indirecta seria únicament una qüestió de prova⁸⁸⁸, sobre la intencionalitat o no de l’empresa d’establir el criteri en base a l’ètnia de les persones que habitaven majoritàriament en el barri. Aquesta posició per part del tribunal va ser criticada perquè semblava tornar a incloure la intencionalitat com a factor a tenir en compte en la normativa antidiscriminació⁸⁸⁹.

3.2.3. Cas Hakelbracht i altres

En aquesta sentència el TJUE eixampla de nou els contorns de la discriminació per associació i ofereix la interpretació més extensa fins al moment d’aquesta figura⁸⁹⁰.

Els fets del cas són els següents: la senyora Hakelbracht va ser entrevistada per a una feina de venedora per la senyora Vandebon, gerent de l’empresa, i li va comunicar que estava embarassada de tres mesos. La senyora Vandebon va informar a l’empresa que la senyora Hakelbracht era una candidata adequada pel lloc de treball i que estava embarassada. L’empresa va comunicar a la senyora Vandebon per correu electrònic la seva negativa de contractar a la senyora Hakelbracht atès que estava embarassada, i la

⁸⁸⁶ TJUE (Gran Sala), sentència de 16 de juliol de 2015, C-83/14, CHEZ Razpredelenie Bulgaria AD contra Komisia za zashtita ot diskriminatsia, par. 128.

⁸⁸⁷ TOBLER, C., *Indirect discrimination under Directives 2000/43 and 2000/78*, op. cit., 2022, pàg. 87.

⁸⁸⁸ BENEDI LAHUERTA, S., “Ethnic discrimination...”, op. cit., 2016, pàg. 807 i 808.

⁸⁸⁹ CAHN, C., “Court of Justice of the EU Rules Collective and Inaccessible Electrical Metres Discriminate against Roma: CHEZ Razpredelenie Bulgaria AD v. Komisia za zashtita ot diskriminatsia (C83/14)”, *European Journal of Migration and Law*, vol. 18, núm. 1, 2016, pàg. 123 i 124.

⁸⁹⁰ MANEIRO VÁZQUEZ, Y., *La discriminación por asociación...*, op. cit., 2022, pàg. 83.

senyora Vandebon va advertir a l'empresa de la il·legalitat d'aquesta decisió. Se li va confirmar amb un segon correu electrònic la negativa de l'empresa, i va comunicar-ho a la senyora Hakelbracht. Aquesta última va posar una reclamació a l'Institut per a la Igualtat entre dones i homes, que va obrir un procediment en el qual l'empresa va negar que la raó per la no contractació hagués estat l'embaràs. També l'empresa, al mateix temps, es va entrevistar amb la senyora Vandebon per recriminar-li haver sigut la causant de la reclamació de la senyora Hakelbracht. Sis mesos després de l'entrevista, se li va comunicar a la senyora Vandebon l'extinció del seu contracte per causes objectives. Les dues van demandar l'empresa a la qual van demanar una indemnització. Davant la negativa de l'empresa, van iniciar un procediment en el Tribunal social d'Amberes que va reconèixer la discriminació directa per raó de sexe patida per la senyora Hakelbracht, però va denegar a la senyora Vandebon la seva condició de víctima d'una represàlia derivada de la seva participació com a testimoni en el cas obert per l'Institut d'Igualtat entre dones i homes, atès que el reconeixement de la condició de testimoni oficial estava subordinada, en la llei belga, al reconeixement de certs requisits formals⁸⁹¹. El TJUE va afirmar que l'article 24 de la Directiva 2006/54/CE no subordina la protecció dels testimonis a requisits formals, sinó que ha d'entendre's en un sentit ampli que inclogui en la seva protecció a totes les persones treballadores que puguin ser objecte de mesures de represàlia per la persona empresària com una reacció a una reclamació presentada per una discriminació basada en el sexe⁸⁹².

Es pot entendre llavors que la protecció pot ser exigida per totes les persones treballadores en sentit ampli, diferents de la persona discriminada per raó de sexe, estenent la protecció a un major nombre de persones, i garantint el mateix nivell i intensitat de tutela. És irrellevant si el suport prestat va ser de manera formal o informal, defensant a la persona o actuant com a testimoni a través d'un procés intern, administratiu o judicial⁸⁹³.

No es requereix una vinculació subjectiva de les persones afectades, de manera que podran ser companyes de feina i pertànyer a la mateixa empresa, o no (s'aplica també aquí la reflexió realitzada a CHEZ, en la qual no es requereix cap connexió entre les persones afectades més enllà de la seva condició de codiscriminades). Tampoc es

⁸⁹¹ TJUE (Sala Tercera), sentència de 20 de juny de 2019, C-404/18, Jamina Hakelbracht y otros contra WTG Retail BVBA, par. 10-21.

⁸⁹² *Ídem*, par. 37.

⁸⁹³ MANEIRO VÁZQUEZ, Y., *La discriminación por asociación...*, op. cit., 2022, pàg. 85 i 86.

requereix la identitat subjectiva entre l'empresari/a causant de la discriminació i aquell que ha pres les represàlies, de manera que un i altre poden ser persones diferents. Els únics requisits imprescindibles serien que la represàlia hagués derivat del suport de la persona treballadora a la víctima de discriminació, i per altra part, que la represàlia provinguí de la persona empresària per a la qual treballa (coincideixi o no amb la de la víctima)⁸⁹⁴.

3.2.4. Assumptes Coleman, CHEZ i Hakelbracht i altres des de la perspectiva de la discriminació algorítmica per associació i per error

Els casos Coleman, Chez i Hakelbracht i altres, amplien l'àmbit personal del dret a la no discriminació a aquelles persones que malgrat no formen part del grup desavantatjat, pateixen discriminació a causa del tracte desfavorable conferit a persones que sí que formen part del grup desavantatjat. La fortalesa de la connexió entre les persones codiscriminades ha anat variant i engrandint els seus límits en cada una de les sentències dictades pel TJUE en la matèria: en el cas Coleman era una connexió entre familiars, en el cas CHEZ va ser evident que no era necessari un vincle familiar o especial relació de cura sinó que era suficient una relació de veïnat, i en el cas Hakelbracht i altres, es va admetre una connexió purament accidental entre les dues víctimes de discriminació derivada de la participació d'una persona en el procés de selecció d'una altra i la seva posició com a testimoni del procés intern per considerar la discriminació⁸⁹⁵. Si el TJUE hagués decidit no garantir la protecció a terceres parts, hagués causat un efecte indesitjable degut a la por a les conseqüències negatives d'associar-se amb persones que ostentin l'atribut protegit (per exemple, persones que no vulguin habitar en àrees on viuen persones d'ètnia gitana)⁸⁹⁶.

A l'hora d'abordar la protecció que les directives ofereixen, interessa destacar l'ús d'un llenguatge uniforme per definir les causes discriminatòries, de manera que per definir la causa determinant del tracte pejoratiu proscriu per les directives, s'empra una fórmula

⁸⁹⁴ *Ídem*, pàg. 89.

⁸⁹⁵ *Ídem*, pàg. 182-183.

⁸⁹⁶ WACHTER, S., "Affinity profiling and discrimination...", *op. cit.*, 2020, pàg. 402.

àmplia, en la qual destaquen les expressions “*per raó de*” o “*per causa de*”. Aquestes fórmules possibiliten tant l’ampliació dels contorns de la causa discriminatòria corresponent, com l’ampliació del número de subjectes emparats sota el seu paraigües de protecció⁸⁹⁷.

Això significa que els casos en què, per exemple, dades de comportament determinessin un atribut protegit no relatiu precisament a la persona en qüestió, sinó a una persona que es relaciona amb ella, es podria aplicar la doctrina de la discriminació per associació⁸⁹⁸.

Per exemple, en el cas dels sistemes algorítmics, la pràctica que poden dur a terme empreses asseguradores d’exigir en determinats barris, en general, unes primes superiors, pot constituir una discriminació en el sentit de l’article 21 de la CDFUE, si la població dels barris pertany a un determinat grup ètnic, classe social o comunitat religiosa. Malgrat alguns dels habitants del barri no pertanyin al grup ètnic, classe social o comunitat religiosa, seran codiscriminats, ja que se’ls aplicaran també les primes més elevades.

La discriminació per associació també pot aplicar-se a la distribució discriminatòria de publicitat que poden generar els sistemes algorítmics, eixamplant el grup de persones potencials a plantejar -i/o a donar suport a- una reclamació de discriminació⁸⁹⁹.

Els casos *Chez*, *Coleman* i *Hakelbracht* i altres, també generen possibilitats per al concepte discriminació per assumpció o per error, ja que admeten la possibilitat que persones que no ostenten pròpiament l’atribut protegit puguin considerar-se víctimes de discriminació⁹⁰⁰.

⁸⁹⁷ MANEIRO VÁZQUEZ, Y., *La discriminación por asociación...*, op. cit., 2022, pàg. 40.

⁸⁹⁸ TJCE (Gran Sala), sentència de 17 de juliol de 2008, C-303/06, S. Coleman contra Attridge Law and Steve Law. En el paràgraf 46 s’exposa que “*aunque en el apartado 56 de la sentencia Chacón Navas, antes citada, el Tribunal de Justicia precisó que, a la vista del tenor literal del artículo 13 CE, el ámbito de aplicación de la Directiva 2000/78 no puede ampliarse a otros tipos de discriminación además de las basadas en los motivos enumerados con carácter exhaustivo en el artículo 1 de la propia Directiva, (...), este Tribunal, sin embargo, no declaró que el principio de igualdad de trato y el alcance ratione personae de dicha Directiva deban interpretarse de manera restrictiva en lo que atañe a los motivos de que se trata*”.

⁸⁹⁹ WACHTER, S., “Affinity profiling and discrimination...”, op. cit., 2020, pàg. 394 i 397. L’autora també esmenta que el concepte de discriminació per associació, al incloure persones que no són part del grup protegit, genera la possibilitat a les persones que sí que formen part del grup protegit però no volen revelar-ho públicament, de presentar una reclamació per discriminació sense significar-se.

⁹⁰⁰ MANEIRO VÁZQUEZ, Y., *La discriminación por asociación...*, op. cit., 2022, pàg. 74.

La discriminació per error⁹⁰¹ es diferencia de la discriminació per associació en què no està involucrada una altra persona que efectivament té una característica protegida, sinó que existeixen certes assumpcions (o adscripcions) que no estan necessàriament basades en fets⁹⁰². Per exemple, en l'assumpte *Accept*⁹⁰³, es va discriminar a un jugador de futbol en base a la seva presumpta orientació sexual i no es va veure obligat a revelar la seva orientació sexual perquè el TJUE pogués valorar la discriminació. És a dir, la mera assumpció d'aquesta característica com a rellevant per prendre una decisió va ser considerada discriminatòria. La Comissió Europea ha expressat que la Directiva 2000/43/EC i la Directiva 2000/78/EC també prohibeixen situacions en les quals una persona és directament discriminada en base a la percepció o assumpció errònia de tenir una característica protegida⁹⁰⁴.

Així i tot, aquesta interpretació no ha estat consistentment aplicada pel TJUE en casos de discriminació. Per exemple, en l'assumpte *Kaltoft*⁹⁰⁵, un home que cuidava infants va ser considerat per qui el contractava com a discapacitat a causa de la seva obesitat, i se'l va acomiadar en part en base a això, ja que es va considerar que aquest factor dificultava la seva feina. El TJUE va ordenar al tribunal nacional que descobrís si aquesta discapacitat existia realment, ja que seria el que determinaria l'existència de discriminació en base a la discapacitat, en comptes de reconèixer que era la percepció d'obesitat com a discapacitat per part de la persona que contractava la raó per la qual hi va haver un tractament diferencial i finalment una discriminació. L'aplicació restrictiva del concepte de discriminació directa en els casos en els quals els motius protegits són percebuts, adscrits o assumits és problemàtica tenint en compte que les tècniques algorítmiques d'elaboració de perfils, que cerquen comportaments, interessos i afinitats, arriben a conclusions sobre les identitats de les persones i grups sota aquesta base. Si la discriminació directa no s'amplia a aquestes situacions, serà difícil de copsar les

⁹⁰¹ La Llei 15/2022 defineix la discriminació per error (també anomenada discriminació per assumpció) en el seu article 6.2.b. També, la Llei 19/2020 defineix aquest concepte en l'article 4, apartat p.

⁹⁰² GERARDS, J. i XENIDIS, R., *Algorithmic Discrimination in Europe...*, op. cit., 2021, pàg. 68.

⁹⁰³ TJUE (Sala Tercera), sentència de 25 d'abril de 2013, C-81/12, *Asociația Accept* contra Consiliul Național pentru Combaterea Discriminării.

⁹⁰⁴ EUROPEAN COMMISSION, *Joint Report on the application of Council Directive 2000/43/EC of 29 June 2000 implementing the principle of equal treatment between persons irrespective of racial or ethnic origin ('Racial Equality Directive') and of Council Directive 2000/78/EC of 27 November 2000 establishing a general framework for equal treatment in employment and occupation ('Employment Equality Directive')*, COM (2014) 2 final, Brussels, 17 January 2014, recital 4.5.

⁹⁰⁵ TJUE (Sala Quarta), sentència de 18 de desembre de 2014, C-354/13, *Fag og Arbejde (FOA) contra Kommunernes Landsforening (KL)*, par. 64.

situacions on les identitats adscrites a través d'algoritmes condueixen a tractament diferencial i desavantatjós de persones o grups⁹⁰⁶.

Wachter posa l'exemple d'un sistema algorítmic que detecta que les persones usuàries de pàgines web sobre Jazz i Blues també veuen anuncis de menjar caribeny. Aquesta decisió no es basa en la raça o l'ètnia sinó en la correlació estadística que suggereix que les persones que tenen interès en certa música també tenen interès en cert tipus de menjar. A partir d'aquí, resultats perjudicials poden ocórrer, si una evidència estadística estableix que un interès en la música Jazz/Blues i en el menjar caribeny es dona més en un grup ètnic en particular i aquest grup és tractat menys favorablement que altres. Totes les persones d'aquest grup, siguin o no del grup ètnic (poden no ser-ho, per exemple, si se'ls ha identificat erròniament com a membre del grup per una classificació inexacta) podran presentar una reclamació de discriminació⁹⁰⁷.

Recapitulant, la jurisprudència del TJUE ha establert l'existència de discriminació directa quan existeixen factors indissociables a l'atribut protegit, possibilitant que l'ús de certs indicadors en els sistemes algorítmics pugui considerar-se discriminació directa. Quan el vincle entre indicador i atribut protegit no sigui suficientment evident, es podria recórrer a la discriminació indirecta, que es basa en els efectes discriminatoris de la mesura. També, arran de la jurisprudència del TJUE, s'ha perfilat la figura de la discriminació per associació, que podia emprar-se, per exemple, en el cas que dades de comportament discriminin en base a un atribut protegit no relatiu precisament a la persona en qüestió, sinó a una persona que es relaciona amb ella, o la figura de la discriminació per error, que podria emprar-se, per exemple, en el cas que dades de comportament identifiquin erròniament, arran d'una classificació inexacta, a una persona com a membre d'un grup que ostenta un atribut protegit. Tanmateix, i com ja hem esmentat, sense accés a informació sobre el sistema o sobre el resultat discriminatori, plantejar prova per tal que es pugui entrar a valorar si la discriminació és directa o indirecta (i dins d'aquestes per associació o per error) pot ser molt difícil. Tot seguit estudiem si el TJUE, a través de la seva jurisprudència, ha obert les possibilitats del plantejament de prova en els casos de discriminació.

⁹⁰⁶GERARDS, J. i XENIDIS, R., *Algorithmic Discrimination in Europe...*, *op. cit.*, 2021, pàg. 68 i 69.

⁹⁰⁷WACHTER, S., "Affinity profiling and discrimination...", *op. cit.*, 2020, pàg. 400 i 401.

4. Possibilitats de millora de la prova

En el quart apartat del capítol II, hem estudiat les dificultats per aportar prova de discriminació algorítmica, a saber, el caràcter marcadament individual de la protecció antidiscriminatòria, la manca d'informació i d'eines d'accés, i la dificultat per l'establiment tant d'un grup desavantajat com d'un grup comparador (necessaris en el cas de discriminació indirecta), a causa de la capacitat dels sistemes algorítmics per excloure a persones directament sense que en siguin conscients, i de discriminar -el més probablement de manera subtil i no evident- en base a múltiples factors -que atomitzin els grups perjudicats i els comparadors.

En aquest apartat, i després d'estudiar en l'anterior les possibilitats que obre la jurisprudència del TJUE establerta en els casos Coleman, Chez, Hakelbracht i altres, Dekker, Tadao Maruko, Jenkins i Bilka-Kaufhaus, volem estudiar les sentències del TJUE que semblen millorar les perspectives per presentar prova en els casos de discriminació algorítmica. Analitzem, doncs, la no necessitat de condició de víctima -cas Feryn-, la suficiència d'un dany hipotètic -cas Accept i Rete Lenford- (4.1), i la possibilitat de considerar com a factor rellevant la denegació d'accés a la informació per invertir de la càrrega de la prova -cas Meister- (4.2).

4.1. No necessitat de condició de víctima (cas Feryn) i suficiència d'un dany hipotètic (cas Accept i Rete Lenford)

En el cas Feryn una persona empresarària va declarar en públic, en el marc d'un procés de contractació, que les sol·licituds de feina de persones de cert origen ètnic no serien admeses. El TJUE va considerar l'existència de discriminació directa⁹⁰⁸.

⁹⁰⁸TJCE (Sala Segona), sentència de 10 de juliol de 2008, C-54/07, Centrum voor gelijkheid van kansen en voor racismebestrijding contra Firma Feryn NV, par. 28.

Per una banda, en l'àmbit processal, en el cas sorgia el dubte sobre si a falta d'una persona demandant identificable que hagués sigut víctima de discriminació, s'aplicava la Directiva 2000/43/CE (la demanda va ser interposada pel Centre belga per a la Igualtat d'Oportunitats i la Lluita contra el racisme). L'article 7 de la Directiva 2000/43/CE exigeix que els estats membres garanteixin l'existència de procediments judicials dels quals puguin valer-se totes les persones que es considerin perjudicades per la no aplicació, en allò que a elles es refereix, del principi d'igualtat de tracte, i les entitats d'interès públic que actuïn en nom de la persona demandant o en el seu suport⁹⁰⁹. No obstant això, l'advocat general d'aquest cas va apuntar que aquesta disposició no impedeix que els estats membres creïn altres disposicions addicionals per obtenir el compliment o la tutela judicial, com per exemple, donar la possibilitat a entitats d'interès públic a exercir una acció judicial també quan no actua en nom d'una persona demandant concreta⁹¹⁰. La Directiva, tanmateix, no obliga els estats membres a assegurar que es reconegui la legitimació de les entitats d'interès públic per exercir accions judicials a falta d'una persona demandant que al·legui haver patit discriminació⁹¹¹ (aquesta qüestió la tractarem amb més detall en el capítol VII).

Per l'altra banda, en relació amb els fets, una declaració pública sobre com les sol·licituds de persones de cert origen ètnic no serien admeses té un efecte que no és hipotètic. En aquests cas, les persones no només quedaran excloses sinó que és molt difícil identificar víctimes individualitzades, ja que les persones afectades pot ser que no sol·licitin la feina. Per tant, en aquesta declaració pública, qui contracta no parla merament de discriminar, sinó que està discriminant. Si es considerés aquest tipus d'accions com excloses del concepte de discriminació directa, perjudicarien -en lloc de promoure- les condicions per a un mercat de treball que propiciï la integració social⁹¹².

En el mateix sentit, en el cas *Accept*, el TJUE va considerar que constituïen fets que permeten presumir l'existència de discriminació les declaracions públiques pronunciades

⁹⁰⁹ Article 7 de la Directiva 2000/43/CE. De la mateixa manera: article 9 de la Directiva 2000/78/CE; article 17 de la Directiva 2006/54/CE i article 8 de la Directiva 2004/113/CE.

⁹¹⁰ A Espanya, l'article 11 bis.2 de la LEC planteja la legitimació per la defensa del dret a la igualtat de tracte i no discriminació quan les persones són una pluralitat indeterminada o de difícil determinació. Aquest article va ser introduït arran de la disposició addicional 5.1 de la Llei Orgànica 3/2007, i ampliat per la disposició final 2.1 de la Llei 15/2022.

⁹¹¹ Conclusiones del abogado general Miguel Poiares Maduro, presentadas el 12 de marzo de 2008, en el asunto, C-54/07, *Centrum voor Gelijkheid van Kansen en voor Racismebestrijding* contra *Firma Feryn NV*, par. 12 i 13.

⁹¹² *Ídem*, par. 15, 16 i 17.

per qui es presentava com a directiu d'un club de futbol professional (i així ho percebia l'opinió pública), en les quals exclouïa de la contractació un futbolista que es presentava com a homosexual⁹¹³. En el cas Rete Lenford, el TJUE va considerar que les declaracions d'un advocat en una entrevista radiofònica en les quals manifestava que no contractaria a una persona homosexual per treballar en el seu despatx d'advocats, malgrat no existir un procés de selecció en curs, també poden entrar en el concepte de "condicions d'accés al treball (...) i a l'exercici professional" continguts en l'article 3.1.a de la Directiva 2000/78/CE, sempre que el vincle entre les declaracions i l'accés al treball i a l'exercici professional de l'empresa no fos hipotètic⁹¹⁴.

Per tant, a Feryn, el TJUE va determinar que dissuadir les persones candidates pertanyents a grups protegits d'optar a un lloc de treball es considerava discriminatori, afegint-se amb els casos Accept i Rete Lenford, que es considerava discriminatori inclús malgrat no existís un procés de selecció en curs -de manera que el TJUE va admetre la suficiència d'un dany hipotètic per establir una discriminació *prima facie*. També, contràriament al TEDH, que considera essencial que existeixi una víctima identificable que s'hagi vist directament afectada per poder presentar la demanda⁹¹⁵, el TJUE va argumentar que l'article 7 de la Directiva 2000/43/CE no impedeix que els estats membres estableixin en la legislació nacional el dret a les associacions amb un interès legítim que compleixin amb els requisits establerts en la directiva⁹¹⁶, o a organismes d'igualtat nacionals (establerts en l'article 13 de la mateixa directiva), d'iniciar procediments legals o administratius per fer complir les obligacions que s'estableixen en la Directiva, sense actuar en nom d'una persona demandant específica o en absència d'una persona demandant identificable⁹¹⁷. Això pot ser de vital importància en cas que una associació o organisme d'igualtat nacional vulgui dur a terme una acció col·lectiva per litigar en relació a l'impacte d'un sistema algorítmic, podent fer-ho en relació al dany hipotètic que està causant (com veurem en el capítol VII i VIII del treball).

⁹¹³TJUE (Sala Tercera), sentència de 25 d'abril de 2013, C-81/12, Asociația Accept contra Consiliul Național pentru Combaterea Discriminării, par. 53.

⁹¹⁴TJUE (Gran Sala), sentència de 23 d'abril de 2020, C-507/18, NH contra Associazione Avvocatura per i diritti LGBTI - Rete Lenford, par. 58.

⁹¹⁵CONSEIL DE L'EUROPE i COUR EUROPÉENNE DES DROITS DE L'HOMME, *Guide pratique sur la recevabilité*, Strasbourg, 31 août 2022, par. 17-20.

⁹¹⁶Vegeu, en el mateix sentit: TJUE (Gran Sala), sentència de 23 d'abril de 2020, C-507/18, NH contra Associazione Avvocatura per i diritti LGBTI - Rete Lenford, par. 65.

⁹¹⁷TJUE, sentència de 10 de juliol de 2008, C-54/07, Centrum voor gelijkheid van kansen en voor racismebestrijding contra Firma Feryn NV, par. 26-27.

Seguint la doctrina d'aquests casos es podrien copsar els danys col·lectius dels biaixos algorítmics que exclouen a persones de l'accés a certs béns i serveis, aplicant-se a la distribució estereotipada de publicitat per part de plataformes d'optimització⁹¹⁸. La falta o la reducció de publicitat a persones integrants de grups protegits podria considerar-se discriminació si, en efecte, soscava l'objectiu de la legislació europea antidiscriminatòria de garantir l'accés igual al mercat de treball⁹¹⁹. En aquest sentit, la personalització de la publicitat que ofereixen els sistemes algorítmics pot ser directament contrària a la regulació antidiscriminatòria, ja que no dona accés a certes ofertes de feina segons les característiques que et defineixen⁹²⁰.

Així i tot, cal puntualitzar que en aquests casos es va considerar el comportament que es descrivia com discriminació directa, atès que hi havia una declaració pública que dissuadia hipotètiques persones candidates a presentar-se a un lloc de treball en aquella empresa concreta. En el cas dels sistemes algorítmics, si es produeix una exclusió, no és pública (sembla difícil que les persones excloses de cert contingut publicitari en siguin conscients), i, per tant, el raonament jurídic aplicat en els casos analitzats pel TJUE podria no ser aplicable en aquest sentit. Sí que ho seria en relació amb l'efecte dissuasiu de l'exclusió, si pensem que després de ser rebutjada sistemàticament sense tenir cap resposta al respecte, la persona pertanyent a un grup protegit pot desistir de presentar-se a posicions de feina on es faci un cribratge de currículums automatitzats.

Tanmateix, perquè aquesta exclusió reiterada pugui ser provada, s'hauria de considerar l'exclusió com a decisió, perquè pogués generar una responsabilitat de notificació. És a dir, en el cas de la intervenció humana significativa que planteja l'article 22.3 de l'RGPD, s'haurien d'incloure en aquest àmbit no només les decisions que puguin considerar-se definitives (per exemple, la decisió de contractació), sinó totes aquelles amb rellevància per a les persones implicades (com l'exclusió dels processos de contractació) que suposaria que a part de què la fase de cribratge fos supervisada per éssers humans, aquests assumissin la responsabilitat de notificar-ho a les persones excloses, facilitant que aquestes poguessin impugnar la decisió⁹²¹.

⁹¹⁸ XENIDIS, R., "Turning EU equality law to algorithmic discrimination...", *op. cit.*, 2020, pàg. 750.

⁹¹⁹ GERARDS, J. i XENIDIS, R., *Algorithmic Discrimination in Europe...*, *op. cit.*, 2021, pàg. 58.

⁹²⁰ XENIDIS, R. i SENDEN, L., "EU Non-discrimination law...", *op. cit.*, 2020, pàg. 166-167.

⁹²¹ RIVAS VALLEJO, P., "Sesgos de género en el uso...", *op. cit.*, 2022, pàg. 75.

4.2. Accés a la informació: Cas Meister

En el cas Meister, una dona russa nascuda el 1961 va demandar a una empresa alemanya, que l'havia rebutjat en dues ocasions per a un lloc de treball, per discriminació per raó de sexe, edat i origen ètnic. El tribunal nacional preguntava al TJUE en primer lloc, si la persona rebutjada tenia dret a ser informada (en base als articles 19.1 de la Directiva 2006/54/CE, 8.1 de la Directiva 2000/43/CE i 10.1 de la Directiva 2000/78/CE⁹²²) sobre si l'empresa havia contractat una altra candidata i en base a quins criteris; i en segon lloc si el fet que la persona empresària no comunicés la informació requerida permetia presumir l'existència de la discriminació al·legada per la persona treballadora.

En relació amb la primera qüestió, els articles que s'enumeren són idèntics a l'article 4.1 de la Directiva 97/80/CE del Consejo, de 15 de diciembre de 1997, relativa a la carga de la prueba en los casos de discriminación por razón de sexo⁹²³ (d'ara en endavant, Directiva 97/80/CE), actualment derogada. El TJUE ja es va pronunciar sobre una qüestió molt similar en el cas Kelly⁹²⁴, en el qual va establir que l'article 4.1 de la Directiva 97/80/CE s'havia d'interpretar en el sentit que no es preveia el dret a un candidat a formació professional, que considera que l'accés a aquesta se li ha denegat a causa de la inobservança del principi d'igualtat de tracte, a accedir a la informació que posseeix l'organitzador de la formació sobre les qualificacions d'altres candidats, per tal que li sigui possible acreditar fets que permetin presumir l'existència de discriminació directa o indirecta conforme a la disposició⁹²⁵. El Tribunal va admetre, tanmateix, que no es podia excloure que la denegació total d'informació per la part demandada, en el context de la prova dels fets, pogués frustrar la possibilitat de les persones d'invocar els seus drets en via jurisdiccional⁹²⁶.

⁹²² Aquests articles es refereixen a la inversió de la càrrega de la prova, per la qual correspon a la part demandada demostrar que no hi ha hagut vulneració del principi d'igualtat de tracte.

⁹²³ Directiva 97/80/CE del Consejo, de 15 de diciembre de 1997, relativa a la carga de la prueba en los casos de discriminación por razón de sexo. DOCE, L 14/6, 20.01.1998.

⁹²⁴ Conclusiones del abogado general Paolo Mengozzi, presentadas el 12 de enero de 2012, en el asunto C-415/10, Galina Meister contra Speech Design Carrier Systems GmbH, par. 10.

⁹²⁵ TJUE (Sala Segona), sentència de 21 de juliol de 2011, C-104/10, Patrick Kelly contra National University of Ireland (University College, Dublin), par. 38.

⁹²⁶ *Ídem*, par. 39.

Constata l'advocat general del cas Meister que, en la hipòtesi que la persona candidata al lloc de feina estigui totalment a mercè de la bona voluntat de la persona empresària en allò relatiu a la informació que fa referència als fets que permetin presumir l'existència d'una discriminació, l'equilibri entre la llibertat de la persona candidata i la persona empresària es pot veure trencat⁹²⁷. Proposa llavors que els articles 19.1 de la Directiva 2006/54/CE, 8.1 de la Directiva 2000/43/CE i 10.1 de la Directiva 2000/78/CE siguin interpretats de manera que l'òrgan jurisdiccional nacional valori l'actitud de la persona empresària, sense limitar-se a valorar la seva mera falta de resposta, sinó, al contrari, situant-la en el context fàctic més ampli en el qual es produeix⁹²⁸. El TJUE, no fa seves les conclusions de l'advocat general i únicament estableix que no es pot excloure que una denegació total de l'accés a la informació per la part demandada pot constituir un dels factors que s'han de tenir en compte en el context de l'acreditació dels fets que permeten presumir l'existència de discriminació directa o indirecta⁹²⁹.

D'una banda, aquesta jurisprudència aplicada als sistemes algorítmics genera certes preocupacions. El fet de no poder accedir a aquesta informació obstaculitzarà la possibilitat d'establir una referència comparativa vàlida. Inclús la no divulgació d'informació pot argumentar-se vàlidament sobre la base de la protecció de la propietat intel·lectual, industrial i de la protecció de dades personals⁹³⁰.

La falta d'informació, a més, no només pot afectar a l'establiment d'un cas *prima facie*, sinó que també afecta les eines dels tribunals d'avaluar el test de necessitat, proporcionalitat i justificació objectiva que cal desenvolupar en la discriminació indirecta⁹³¹, essent difícil llavors, avaluar, per exemple, si una empresa de software responsable de la comercialització d'un algoritme discriminatori podria haver optat per

⁹²⁷ Conclusiones del abogado general Paolo Mengozzi, presentadas el 12 de enero de 2012, en el asunto C-415/10, Galina Meister contra Speech Design Carrier Systems GmbH, par. 33.

⁹²⁸ *Ídem*, par. 37.

⁹²⁹ TJUE (Sala Segona), sentència de 19 d'abril de 2012, C-415/10, Galina Meister contra Speech Design Carrier Systems GmbH, par. 47.

⁹³⁰ El Considerant 63 de l'RGPD estableix que el dret d'accés de les persones a les dades personals recollides que li concerneixen no ha d'afectar negativament als drets i llibertats de tercers, inclosos els secrets comercials o la propietat intel·lectual, i en particular, als drets de la propietat intel·lectual que protegeixen programes informàtics. S'estableix també, que malgrat aquestes consideracions, no han de tenir com a resultat la negativa a prestar tota la informació a la persona interessada.

⁹³¹ TJUE (Sala Segona), sentència de 19 d'abril de 2012, C-415/10, Galina Meister contra Speech Design Carrier Systems GmbH. En aquest sentit, vegeu també: TJCE (Sala Segona), sentència de 10 de juliol de 2008, C-54/07, Centrum voor gelijkheid van kansen en voor racismebestrijding contra Firma Feryn NV; TJUE (Sala Tercera), sentència de 25 d'abril de 2013, C-81/12, Asociația Accept contra Consiliul Național pentru Combaterea Discriminării; TJUE (Gran Sala), sentència de 23 d'abril de 2020, C 507/18, NH contra Associazione Avvocatura per i diritti LGBTI - Rete Lenford.

una solució menys discriminatòria tenint en compte les explicacions altament tècniques que probablement l'empresa presenti⁹³². En resum, la negativa a revelar informació difícilment podrà ser emprada com a element indiciari per provar l'existència de discriminació⁹³³.

D'altra banda, aquesta jurisprudència reconeix la dificultat de provar un cas de discriminació sense tenir informació, és més, reconeix que la denegació total d'informació pot frustrar les possibilitats de les persones d'invocar els seus drets. Donades les conseqüències que suposa l'asimetria d'informació en aquests casos, la Comissió va impulsar l'any 1988 la Propuesta de Directiva del Consejo relativa a la carga de la prueba en el ámbito de la igualdad de retribución y trato entre hombres y mujeres (sense tenir en compte l'existència de sistemes algorítmics que agreujarien la situació) en la qual plantejava la consagració d'un dret a la informació en favor de les víctimes de discriminació. L'article 4.b de la proposta estava redactat de la següent manera: "*las partes en litigio puedan disponer de cualquier información pertinente en poder de la otra parte, o que ésta pueda razonablemente obtener, que sea necesaria para hacer valer sus derechos y cuya divulgación no lesione gravemente los intereses de la otra parte por motivos distintos de los del litigio*"⁹³⁴. La disposició mai va ser recollida en els texts finals⁹³⁵, però actualment recobra rellevància per les possibilitats que podria donar en un plet de discriminació algorítmica. Anant més enllà, es podria plantejar si en el cas de la discriminació algorítmica, donada l'asimetria d'informació i de poder, s'hauria d'automatitzar una inversió de la càrrega de la prova quan la persona no pugui aportar informació suficientment exhaustiva per constituir un indici sobre els criteris del tribunal -assimilant-ho a la negativa de l'empresa a donar informació-, ja que la persona que s'ha vist discriminada probablement es trobarà sense suficient informació (i si la té, sense eines per comprendre-la) per establir un cas *prima facie*. Aquesta seria una proposta que parteix de la denominada doctrina de la discriminació *per se* dels sistemes algorítmics amb efecte advers sobre les categories protegides, i en la qual es traslladaria a aquella persona que

⁹³² GERARDS, J. i XENIDIS, R., *Algorithmic Discrimination in Europe...*, op. cit., 2021, pàg. 75.

⁹³³ SORIANO ARNANZ, A., "Discriminación algorítmica: garantías...", op. cit., 2022, pàg. 154.

⁹³⁴ COMISIÓN DE LAS COMUNIDADES EUROPEAS, *Propuesta de Directiva del Consejo relativa a la carga de la prueba en el ámbito de la igualdad de retribución y trato entre hombres y mujeres*, (COM (88) 269 final), Bruselas, 27 de mayo de 1988, art. 4.b.

⁹³⁵ Conclusiones del abogado general Paolo Mengozzi, presentadas el 12 de enero de 2012, en el asunto C-415/10, Galina Meister contra Speech Design Carrier Systems GmbH, par. 21.

pren la decisió la càrrega de provar que la decisió no és discriminatòria o que és necessària malgrat que és discriminatòria⁹³⁶.

A tall de recapitulació, el TJUE constata en el cas Feryn que no es necessària una víctima identificada per portar un cas de discriminació davant del tribunal. En el cas *Accept i Rete Lenfod* estableix, a més, la suficiència d'un dany hipotètic. Tanmateix, volem plantejar que malgrat la doctrina s'hagi emparat en aquests casos per obrir les possibilitats en els casos de discriminació algorítmica, es tractava de casos molt concrets en els quals hi havia una declaració pública que evidenciava com un grup determinat de població que compartia una característica protegida s'havia vist discriminat. En el cas de la discriminació algorítmica, no hi haurà unes declaracions que dissuadeixin sinó que directament hi haurà exclusió, i llavors, el raonament jurídic del TJUE podria no ser aplicable. En el cas *Meister*, el TJUE estableix que la denegació total de l'accés a la informació per la part demandada pot constituir un dels factors que s'han de tenir en compte en la determinació d'un cas *prima facie*. Tanmateix, en el cas de la discriminació algorítmica pot ser necessari anar més enllà, i hem destacat que pren especial rellevància l'article 4.b de la Propuesta de Directiva del Consejo relativa a la carga de la prueba en el ámbito de la igualdad de retribución y trato entre hombres y mujeres, on s'establí que les parts en litigi podien disposar de qualsevol informació necessària en poder de l'altra part per fer valer els seus drets⁹³⁷.

5. Conclusiones parciales

En aquest capítol ens proposàvem respondre, en part, a la tercera qüestió de recerca que plantejava si algunes garanties de la regulació antidiscriminació i de protecció de dades podrien ser vàlides per protegir les persones de la discriminació algorítmica, donades certes condicions. Hem plantejat els punts forts que presenta el marc regulador europeu antidiscriminatori i si aquests, sota certes condicions, poden mitigar els punts febles que

⁹³⁶AJUNWA, I., "The paradox of automation as anti-bias intervention", *Cardozo Law Review*, vol. 40, 2020, pàg. 1726 i s. Cal puntualitzar que la doctrina de discriminació *per se* dels sistemes algorítmics s'ha proposat és en base al títol VII de la Civil Rights Act, no en base al sistema europeu.

⁹³⁷COMISIÓN DE LAS COMUNIDADES EUROPEAS, *Propuesta de Directiva del Consejo relativa a la carga de la prueba en el ámbito de la igualdad...*, *op. cit.*, 27 de mayo de 1988, art. 4.b.

hem plantejat en el capítol II o directament ser útils per a l'abordatge de la discriminació algorítmica.

En primer lloc, per mitigar la llista exhaustiva d'atributs protegits que presenta l'article 19 del TFUE i les limitacions de les directives antidiscriminació, s'han exposat els avantatges de proporcionar un esquema de protecció més ampli, que operi amb independència d'atributs protegits específics, abraçant la redacció de provisions com l'article 14 del CEDH, que prohibeix la discriminació en base a qualsevol atribut, o l'article 21 de la CDFUE, que prohibeix tota discriminació. En aquestes provisions, s'han d'interpretar, llavors, els atributs enumerats com a exemples, donant la possibilitat a organitzacions i tribunals de comparar-los amb altres atributs o indicadors que siguin similars perquè reflecteixin biaixos, prejudicis o estereotips de manera comparable. Quant a les directives antidiscriminació, s'ha subratllat la necessitat d'equiparar la protecció que es brinda a tots els atributs protegits pel que fa als àmbits materials (treball, accés a béns i serveis i sistema del benestar).

En segon lloc, per pal·liar el fet que la normativa antidiscriminació no atorgui suficient rellevància a l'element grupal definitori no eliminable de la discriminació, s'ha plantejat la necessitat de considerar en l'anàlisi de casos concrets per part de tribunals, la discriminació estructural i la discriminació interseccional. Quant a la primera, pot ajudar a interpretar el tracte que produeixi una discriminació a la llum de la posició social estructural de subordinació, no obstant que es manifesti de forma individualitzada; a entendre que els algorismes repliquen i perpetuen les estructures de dominació; i que la perspectiva de la discriminació com a greuge individual que presenta majoritàriament la normativa antidiscriminació no és suficient, essent necessària la instrumentalització del dret com a eina transformadora de les estructures socials creadores dels sistemes d'opressió amb una visió centrada en la lògica subjacent a l'ús de la intel·ligència artificial. Quant a la segona, pot ajudar a millorar la interpretació del resultat d'una decisió presa mitjançant un algorisme, que en molts casos, no basarà la discriminació en un atribut protegit, sinó que ho farà en base a múltiples atributs protegits, possibilitant que es tingui en compte que si determinades decisions automatitzades no comporten un tractament menys favorable en relació amb un atribut protegit, potser sí que el comporten en base a diversos atributs protegits combinats. Hem proposat, com a mesura, la prohibició de l'ús de sistemes d'IA en contexts en els quals no es pugui assegurar que no s'estan perpetuant sistemes d'opressió i subordinació. També, la regulació podria, de manera més pràctica,

assegurar que si més no, la discriminació que produeixen els sistemes algorítmics no és més exacerbada que aquella que produeixen les persones. Per exemple, si l'algoritme no és un 95% precís per a tots els grups de població, tenint en compte també la interseccionalitat (p. e. un grup de població que siguin dones, afroamericanes de més de 60 anys), no es podrà introduir en el mercat. En el mateix sentit, però establint un altre sistema de cribratge, la normativa pot establir que no es poden emprar sistemes en els quals les taxes d'error per tots els grups (incloent categories protegides combinades) superin l'1%.

En tercer lloc, com a punt fort, la jurisprudència del TJUE ha establert l'existència de discriminació directa quan existeixen factors indissociables a l'atribut protegit (p. e. cas Dekker i Tadao Maruko), possibilitant que l'ús de certs indicadors en els sistemes algorítmics pugui considerar-se discriminació directa. En el cas que el vincle entre indicador i atribut protegit no sigui suficientment evident, es podrà recórrer a la discriminació indirecta, que es basa en els efectes discriminatoris de la mesura (p. e. cas Jenkins i Bilka-Kaufhaus). També, arran de la jurisprudència del TJUE s'ha perfilat la figura de la discriminació per associació (p. e. cas Coleman, CHEZ i Hakebrandt i altres) que podrà emprar-se, per exemple, en el cas que dades de comportament discriminin en base a un atribut protegit no relatiu precisament a la persona en qüestió, sinó a una persona que es relaciona amb ella, o la figura de la discriminació per error, que podrà emprar-se, per exemple, en el cas que dades de comportament identifiquin erròniament i perjudiquin, arran d'una classificació inexacta, a una persona com a membre d'un grup que ostenta un atribut protegit. En aquests casos, en haver més persones perjudicades, també hi haurà més persones que puguin desafiar la decisió. Tanmateix, malgrat que hi hagi diferents tipologies de discriminació desenvolupades per la jurisprudència de TJUE on pugui recaure la discriminació algorítmica, si el dret antidiscriminatori no proporciona ajuda per plantejar la prova que aquesta discriminació s'ha produït, no seran útils, fet que ens porta al següent punt.

En quart lloc, el TJUE en el cas Feryn estableix que no és necessària una víctima identificada per portar un cas de discriminació davant del tribunal, i en el cas *Accept i Rete Lenfod*, que és suficient un dany hipotètic, ampliant així les possibilitats de portar un cas de discriminació algorítmica davant del tribunal. Tanmateix, cal tenir en compte que es tractava de casos molt concrets en els quals hi havia una declaració pública que evidenciava com un grup determinat de població, que compartia una característica

protegida, s'havia vist discriminat, i en el cas de la discriminació algorítmica, no hi haurà unes declaracions que dissuadeixin sinó que directament hi haurà exclusió, i llavors, el raonament jurídic del TJUE podria no ser aplicable. En el cas Meister, el TJUE estableix que la denegació total de l'accés a la informació per la part demandada pot constituir un dels factors que s'han de tenir en compte en la determinació d'un cas *prima facie*. No obstant això, en el cas de la discriminació algorítmica pot ser necessari anar més enllà, i hem destacat que pren especial rellevància l'article 4.b de la Propuesta de Directiva del Consejo relativa a la carga de la prueba en el ámbito de la igualdad de retribución y trato entre hombres y mujeres, on s'establí que les parts en litigi podien disposar de qualsevol informació necessària per fer valer els seus drets en poder de l'altra part⁹³⁸.

Per tant, conclouem que malgrat la normativa antidiscriminació europea pugui mitigar, sota certes condicions, els punts febles que presenta, i oferir possibilitats per iniciar casos de discriminació algorítmica davant d'un tribunal, hi ha elements definitoris de la seva estructura que no faciliten la lluita enfront la discriminació algorítmica -i inclús la discriminació en general-, com per exemple, la manca de consideració al component estructural i interseccional de la discriminació o la manca d'eines que facilitin l'accés a prova per part de les persones que s'han vist discriminades. Tot seguit abordem la mateixa qüestió però en relació amb la normativa de protecció de dades europea.

⁹³⁸ *Ídem*.

Capítol V. Aportacions de l'RGPD a la protecció contra la discriminació algorítmica

Havent estudiat en el capítol IV els punts forts que presenta el marc regulador europeu antidiscriminatori i si aquests, sota certes condicions, poden mitigar els punts febles que hem plantejat en el capítol II, o directament ser útils per l'abordatge de la discriminació algorítmica, en aquest capítol desenvolupem la segona part de la tercera qüestió de recerca, a saber, els elements en la normativa en matèria de protecció de dades que creiem que poden tenir incidència positiva en la regulació i disminució de la discriminació algorítmica, i que es contraposen amb les problemàtiques que hem plantejat en el capítol III, seguint el mateix ordre.

En primer lloc, abordem l'RGPD com a normativa basada en la governança col·laborativa i les possibilitats que proporciona en la disminució de la discriminació algorítmica - donant una interpretació alternativa a la funcionalitat dels drets individuals que proporciona l'RGPD-, com per exemple, l'article 22 (1); en segon lloc, les exigències de qualitat de les dades que promou l'RGPD, com a salvaguardes en referència als biaixos algorítmics (2); en tercer lloc, havent estudiat les limitacions dels requisits d'accés a informació i transparència en l'RGPD sota quines condicions poden participar en l'abordatge i lluita envers la discriminació algorítmica (3); en quart lloc, quines poden ser les aportacions de les eines de responsabilitat proactiva i gestió del risc, tal com estan plantejades ara en l'RGPD i estudiant també la possibilitat d'anar més enllà de la regulació actual (4), i finalment, en contraposició al sistema de protecció contra la discriminació a través de drets individuals de l'RGPD, que com hem vist en l'apartat 5 del capítol III, presenta certes limitacions, les possibilitats que planteja l'article 80.2 de l'RGPD, que proporciona la possibilitat a entitats, associacions i organitzacions per defensar drets col·lectius i promoure accions de cessació de tractaments de dades que vulnerin les obligacions en matèria de protecció de dades, integrant l'element grupal necessari per poder abordar la discriminació algorítmica (5). Finalment, resumim els elements tractats en aquest capítol en un apartat de conclusions parcials (6).

1. Possibilitats de l'RGPD com a sistema de governança col·laborativa per abordar la discriminació algorítmica

En aquest apartat, que inicia un capítol sobre l'RGPD com a normativa reguladora de la discriminació algorítmica, estudiem la governança col·laborativa, explorant, en primer lloc, el seu funcionament i les seves implicacions en l'abordatge de la discriminació algorítmica (1.1), i en segon lloc, les condicions requerides perquè esdevingui un règim de protecció efectiu envers la discriminació algorítmica (1.2).

1.1. L'RGPD com a sistema de governança col·laborativa

L'RGPD com a sistema de governança col·laborativa -que es descriu com l'associació pública i privada en la governança⁹³⁹-, crea un sistema de drets individuals combinat amb un sistema de responsabilitat proactiva⁹⁴⁰, on qui té la responsabilitat de regular treballa amb indústries regulades i terceres parts per crear el contingut de la regulació i executar

⁹³⁹ Vegeu articles doctrinals que aposten per la governança col·laborativa per abordar la regulació dels sistemes algorítmics: CITRON, D. K., “Technological Due Process”, *op. cit.*, 2008, pàg. 1308-1312 (proposa combinar la notificació i l'oportunitat de ser escoltat/da amb sistemes d'avaluació o juntes de persones expertes); ANANNY, M., i CRAWFORD, K., “Seeing without Knowing: Limitations...”, *op. cit.*, 2018, pàg. 983 i 984 (exposen les limitacions de la transparència com a únic element regulador dels sistemes algorítmics i aposten per la rendició de comptes); DESAI, D. R. i KROLL, J. A., “Trust but Verify: A Guide to Algorithms and the Law”, *Harvard Journal of Law and Technology*, vol. 31, 2017, pàg. 63 i 64 (els autors defensen que la transparència no pot ser l'únic element regulador dels sistemes algorítmics i s'han d'introduir avaluacions, auditories i supervisions de les autoritats competents); NICHOLSON PRICE II, W., “Regulating Black-Box Medicine”, *Michigan Law Review*, vol. 116, núm. 3, 2017, pàg. 474 (l'autor planteja, per regular els algoritmes mèdics, flexibilitzar el control de la Food and Drug Administration (FDA) i emprar la governança algorítmica); KATYAL, S. K., “Private Accountability in the Age of Artificial Intelligence”, *op. cit.*, 2019, pàg. 61 (l'autora alerta que s'està adreçant l'atenció al lloc equivocat si es mira únicament a l'estat per abordar problemes de rendició de comptes en relació amb els algoritmes).

⁹⁴⁰ KAMINSKI, M. E. i MALGIERI, G., “Algorithmic impact assessments...”, *op. cit.*, 2021, pàg. 126 i 127. En el context de la governança algorítmica, l'RGPD conté drets de protecció de dades generals (drets de notificació, d'accés, de rectificació, de supressió i de limitació del tractament) que també s'apliquen a les decisions automatitzades. També estableix drets específics a la presa de decisions mitjançant algoritmes, que inclouen el dret de notificació d'una decisió únicament automatitzada (art. 13 i 14), el dret de notificació i accés a informació significativa sobre la lògica involucrada (art. 13, 14 i 15); el dret de la persona a ser informada del significat dels efectes de la decisió únicament automatitzada (art. 13, 14 i 15), i el dret a no ser subjecte d'una decisió únicament automatitzada (art. 22) amb salvaguardes i restriccions que limiten els casos on les decisions únicament automatitzades estan permeses, com són el dret a oposar-se a la decisió, a expressar el punt de vista i a la intervenció humana. Si no s'invoquen cap d'aquests drets, les persones responsables de dades igualment han de complir certes obligacions (tant en el cas que hi hagi intervenció humana com no), incloent auditories de terceres parts, la designació d'una persona responsable de protecció de dades (art. 37 RGPD), o la realització d'avaluacions d'impacte (art. 35 RGPD).

les regulacions, sense ser equivalent a l'autoregulació, atès que l'administració pública ostenta un rol central⁹⁴¹.

La governança col·laborativa es basa en la següent premissa: qui empra el tractament de dades personals tem demandes legals i investigacions creades per queixes i l'exercici de drets individuals. Aquesta situació pot generar predisposició a unir-se amb persones reguladores per ajudar a definir què és factible i què no en un sector particular. Per exemple, una companyia que tem rebre una sanció per haver errat en posar en pràctica salvaguardes adequades per evitar la discriminació en la presa de decisions mitjançant algoritmes, és més procliu a negociar un codi de conducta o posar en pràctica processos de compliment normatiu i infraestructura (per exemple, auditories, delegats/des de protecció de dades, avaluacions d'impacte...) per operar d'acord amb les guies. Tanmateix, el sistema admet que les organitzacions decideixin no participar en la governança col·laborativa (que té un cost econòmic i de destinació de recursos) i fer front directament a reptes legals individuals -que suposaran un cost igualment⁹⁴².

En referència a enfrontar la discriminació algorítmica mitjançant els mecanismes plantejats per l'RGPD, d'una banda, el sistema de drets individuals pot afavorir les persones a abordar una problemàtica individual, però sense corregir problemes sistèmics, i d'altra banda, la responsabilitat proactiva promou eines per tractar problemes sistèmics, per exemple amb eines com les auditories, les avaluacions d'impacte, els codis de conducta, certificacions, marques i segells o la supervisió per part d'autoritats independents o grups de persones expertes, però sense entrar a valorar assumptes individuals.

Una persona que és víctima de discriminació podrà fer front individualment a la situació però sense corregir el que ha causat precisament la discriminació. Inclús quan es fa amb bona fe, la correcció individual (o la supressió de certes dades) pot provocar que el sistema de presa de decisió en general sigui menys acurat. El mateix pot passar en relació amb el dret a la intervenció humana en el cas de les decisions preses mitjançant algoritmes, com en el cas que afegint intervenció humana es respecti més la dignitat individual fent que la precisió del sistema en general sigui pitjor, i consegüentment

⁹⁴¹ KAMINSKI, M. E. i MALGIERI, G., "Algorithmic impact assessments...", *op. cit.*, 2021, pàg. 127.

⁹⁴² KAMINSKI, M.E., "Binary Governance...", *op. cit.*, 2019, pàg. 1576 i 1581.

impactant altres individus subjectes a l'algoritme⁹⁴³. De manera que els drets individuals són una eina limitada per trobar i corregir problemes sistèmics com l'error, el biaix i la discriminació dels algoritmes. Això pot ser, o bé a causa de la capacitat limitada de les persones detentores dels drets, ja sigui tècnica, legal o econòmica; o bé ja que descobrir problemes sistèmics com la discriminació seria més fàcil de fer amb una perspectiva general del comportament del sistema, més que avaluant decisions precises; o bé finalment, a causa del temps, ja que en existir algoritmes que evolucionen en el temps, intentar regular la seva qualitat a través de desafiaments individuals és inviable⁹⁴⁴.

Per exemple, l'article 22 de l'RGPD, com hem analitzat en l'apartat 1 del capítol III, no sembla ser suficientment efectiu per abordar la discriminació algorítmica. Si bé en el cas de disposar d'informació sobre la discriminació i entrar dins les condicions que estipula l'article (decisió únicament automatitzada amb efectes significatius o d'afectació similar) es pot aconseguir una correcció individual, no corregeix el sistema en si, que podrà seguir causant discriminació cap a altres individus. En canvi, si s'entén aquest article com a element dissuasiu d'un sistema més ampli i complex que integra mesures sistèmiques, pot participar a generar una política de compliment.

Malgrat que un sistema de governança col·laborativa obre les possibilitats reguladores i d'interpretació de la norma, és necessari preguntar-se per les implicacions *de facto* que comporta, com per exemple que moltes de les propostes actuals deleguen la governança de l'error algorítmic, el biaix i la discriminació a les organitzacions privades que acaben tenint un poder significatiu que no està regulat pel govern⁹⁴⁵, generant dues problemàtiques: en primer lloc, la incidència de les organitzacions privades en la regulació i, en segon lloc, la rendició de comptes que se'ls exigeix.

1.2. Condicions per l'efectivitat d'un règim de governança col·laborativa

En relació amb la incidència de les organitzacions privades en la regulació, en l'àmbit de la intel·ligència artificial, s'ha instaurat un tipus d'autoregulació on la influència de

⁹⁴³ *Ídem*, pàg. 1614.

⁹⁴⁴ *Ídem*, pàg. 1558.

⁹⁴⁵ KATYAL, S. K., "Private Accountability in the Age of Artificial Intelligence", *op. cit.*, 2019, pàg. 106.

companys corporatives guia el comportament dels participants en la indústria⁹⁴⁶. Solow-Niederman alerta que les condicions necessàries per una governança publicoprivada no existeixen actualment i són un objectiu futur molt idealista. Es requereix, per definició, que els dos components siguin iguals, i quan parlem de la intel·ligència artificial, no ho són. Inclús, quan existeixen entitats públiques relatives a la intel·ligència artificial, necessiten dependre del suport privat tant per assistència tècnica com en consells per regular⁹⁴⁷.

El sector privat supera al sector públic tant en relació amb l'expertesa com en relació amb els recursos (els actors privats quan prenen decisions tècniques sobre la intel·ligència artificial també estan prenent decisions polítiques, és a dir, donat el control privat del disseny i el desenvolupament de les tecnologies de la intel·ligència artificial, les decisions preses per actors privats en realitat estan regulant el comportament de les persones) resultant d'això una governança privada *de facto*⁹⁴⁸. Per exemple, la governança col·laborativa sovint suposa requeriments textuais amplis i ambigus, que es clarifiquen durant el temps a través de guies reguladores o millors pràctiques de la indústria o negociació entre empreses i qui regula⁹⁴⁹: implementar “*mesures adequades*”⁹⁵⁰ pot significar una amalgama de normes d'obligat compliment, guies i pràctiques internes de les empreses⁹⁵¹ (pensem, per exemple, en les “mesures adequades” de l'article 22.3); o utilitzar (la traducció és nostra) “(...) *procediments matemàtics o estadístics adequats per a l'elaboració de perfils, aplicar mesures tècniques i organitzatives apropiades per garantir, en particular, que es corregeixen els factors que introdueixen inexactituds en les*

⁹⁴⁶ GUIHOT, M. *et al.*, “Nudging Robots: Innovative Solutions to Regulate Artificial Intelligence”, *Vanderbilt Journal of Entertainment and Technology Law*, vol. 20, núm. 2, 2017, pàg. 427-436. Els autors defensen que el rol tradicional del govern s'ha vist disminuït amb el temps i han proliferat els sistemes d'autoregulació en el buit del control governamental. Els governs que abans tenien una posició central i influència han cedit una part d'aquesta influència a un grup dissipat de participants reguladors. El poder governamental segueix existint, l'autoregulació evoluciona i esdevé la posició per defecte. Clàssicament l'autoregulació tendeix a usurpar el rol de la regulació pública, amb els riscos que això comporta, concretament, que els principis i codis de la indústria no són obligatoris, que manquen règims efectius d'execució i que al suggerir-se moltes perspectives diferents el contingut i les prioritats varien i falten organitzacions governamentals per coordinar la direcció i el compliment.

⁹⁴⁷ SOLOW-NIEDERMAN, A., “Administering Artificial Intelligence”, *Southern California Law Review*, vol. 93, 2020, pàg. 679-684. L'autora comenta que en relació amb la intel·ligència artificial, no hi ha un organisme estatal amb qui forjar una associació públic-privada. No obstant que es pugui necessitar que s'involucri el sector privat per aportar la seva expertesa i perspectiva, sense una veu pública que compensi la visió privada a la qual el públic general pugui retre comptes, aquesta combinació de forces no és un model de negociació públic-privat creïble.

⁹⁴⁸ *Idem*, pàg. 641-644.

⁹⁴⁹ KAMINSKI, M. E. i MALGIERI, G., “Algorithmic impact assessments...”, *op. cit.*, 2021, pàg. 126.

⁹⁵⁰ Per exemple, entre d'altres, en l'article 14, 22, 24 de l'RGPD.

⁹⁵¹ KAMINSKI, M.E., “Binary Governance...”, *op. cit.*, 2019, pàg. 1602.

dades personals i es redueix al màxim el risc d'error (...)”⁹⁵², pot significar que la naturalesa precisa dels procediments i les mesures serà determinada pel sector privat o potser a través d'estàndards tècnics⁹⁵³. I aquí es troba la qüestió principal: l'RGPD, ideat per protegir el drets fonamentals, pretén fer-ho en gran part amb la col·laboració de les organitzacions privades (que no són guardianes de drets i que els seus interessos no s'alineen amb els interessos de les persones individuals)⁹⁵⁴, de manera que la privatització de la protecció de drets fonamentals és prevalent a tot l'RGPD⁹⁵⁵.

Un règim de col·laboració de bona fe amb el sector privat té un fort potencial de ser soscavat pels incentius i les lògiques del sector privat⁹⁵⁶. Una visió crítica de la governança col·laborativa defensa que les empreses no incorporen simplement els requisits legals en les seves pràctiques sinó que el procés de fer-ho deforma els requisits legals perquè encaixin en la lògica organitzacional existent de l'empresa⁹⁵⁷.

Sorgeix la qüestió de si el TJUE trobarà que la perspectiva de governança col·laborativa de l'RGPD és adequadament protectora dels drets fonamentals. Lee examina com arran de la sentència del TJUE, *Google Spain contra Agencia Española de Protección de Datos*⁹⁵⁸, el tribunal ha posicionat Google com una agència administrativa privada⁹⁵⁹,

⁹⁵² Considerant 71 de l'RGPD. La perspectiva de l'RGPD per prevenir els biaixos i les discriminacions en les decisions preses mitjançant algoritmes és un altre exemple de governança col·laborativa.

⁹⁵³ KAMINSKI, M.E., “Binary Governance...”, *op. cit.*, 2019, pàg. 1603.

⁹⁵⁴ SELBST, A. D., “An institutional view of Algorithmic Impact Assessments”, *Harvard Journal of Law and Technology*, vol. 35, núm.1, 2021, pàg. 168. L'autor defensa, que per les empreses tecnològiques, l'ètica està fora de la seva competència. Poden contractar persones expertes en ètica, però fonamentalment les estructures organitzacionals, actituds, i experiències estan fetes per construir tecnologia per treure'n profit. Vegeu, també, en aquest sentit: BIETTI, E., “From Ethics Washing to Ethics Bashing: A Moral Philosophy View on Tech Ethics”, *Journal of Social Computing*, vol. 2, núm. 3, 2021, pàg. 267. L'autor defensa que l'ús corporatiu del llenguatge ètic en cercles tecnològics s'ha usat per part de les companyies com una façana que justifica la desregulació, autoregulació i la governança dominada pel mercat. Aquesta instrumentalització s'ha anomenat “ethics washing”.

⁹⁵⁵ KAMINSKI, M.E., “Binary Governance...”, *op. cit.*, 2019, pàg. 1615.

⁹⁵⁶ SELBST, A. D., “An institutional view of Algorithmic Impact Assessments”, *op. cit.*, 2021, pàg. 117. L'autor utilitza el terme “legal managerialism” observant com donar confiança a empreses en relació amb la implementació de regulació causa la corrupció de la regulació per les lògiques del mercat que dominen les firmes, i que el fet que treballi conjuntament aquell qui regula i qui és regulat pot minar l'objectiu regulador. Especifica també que quan s'abraça el sector privat com a col·laborador, la rendició de comptes passa de l'obligació d'execució a l'encoratjament.

⁹⁵⁷ EDELMAN, L. B., *Working Law: Courts, Corporations, and Symbolic Civil Rights*, Chicago, University of Chicago Press, 2016, pàg. 142-146.

⁹⁵⁸ TJUE (Gran Sala), sentència de 13 de maig de 2014, C-131/12, *Google Spain S.L. contra Agencia Española de Protección de Datos*.

⁹⁵⁹ LEE, E., “Recognizing Rights in Real Time: The Role of Google in the EU Right To Be Forgotten”, *University of California, Davis Law Review*, vol. 49, núm. 3, 2016, pàg. 1025.

acceptant la participació de l'empresa en la determinació del contingut del dret a l'oblit implementat en un context específic⁹⁶⁰.

En relació amb la rendició de comptes que s'ha d'exigir a les organitzacions privades, creiem necessari algun tipus de mecanisme per fomentar o requerir les persones desenvolupadores de la intel·ligència artificial que incorporin funcions de seguretat en els seus sistemes de manera que es necessitarà establir clarament la responsabilitat pels tipus d'errors o problemes sistèmics que es volen evitar⁹⁶¹. Nogensmenys, s'ha criticat que el caràcter obert i abstracte de les normes de l'RGPD –per tal que siguin adaptables als canvis constants de la tecnologia-, ocasiona, no només que les normes siguin difícils d'aplicar⁹⁶², sinó que sigui també difícil el control de la seva aplicació. Soriano Aranz planteja que principalment la problemàtica de la regulació de protecció de dades per abordar els riscos de la discriminació algorítmica i oferir salvaguardes, rendició de comptes i solucions rau en què a part que el seu objectiu no és ser un instrument antidiscriminatori, els seus articles a la pràctica han esdevingut dret tou⁹⁶³.

Per crear un règim de governança efectiu o legítim, les eines de responsabilitat proactiva han d'anar acompanyades de sistemes supervisors⁹⁶⁴ (per exemple, auditories amb una amenaça executiva si no es duen a terme, avaluacions d'impacte amb transparència pública, o òrgans supervisors amb l'habilitat d'executar sancions o publicar resultats⁹⁶⁵). Sense regulació pública o supervisió, els actors comercials no estaran necessàriament incentivats per respondre més directament a les persones, sobretot si s'és d'un col·lectiu el qual el sistema d'intel·ligència artificial sistemàticament subestima. Solow-Niederman cita un cas de la Universitat de Medicina de Chicago on, en el cas d'haver-se aplicat, l'algoritme hagués guiat cap al resultat paradoxal d'aplicar més recursos per a la població més privilegiada en comptes de la població més vulnerable. L'actor privat pot veure aquestes problemàtiques en certes ocasions, ja sigui per un desig de fer les coses bé o per evitar l'oprobri, però sense una forma més consistent de revisió pública no hi ha garanties que aquest tipus de governança protegeixi els drets i llibertats, hi hagi equilibri entre riscos

⁹⁶⁰ KAMINSKI, M.E., "Binary Governance...", *op. cit.*, 2019, pàg. 1615.

⁹⁶¹ *Ídem*, pàg. 1573.

⁹⁶² ZUIDERVEEN, F., *Discrimination, artificial intelligence...*, *op. cit.*, 2018, pàg. 45.

⁹⁶³ En anglès s'anomena "soft law".

⁹⁶⁴ Com a mecanismes per assegurar l'aplicació de l'RGPD es compta amb la designació de delegats de protecció de dades (article 37-39 RGPD), el Comitè Europeu de Protecció de Dades (article 68 RGPD), les Autoritats de Protecció de Dades dels estats membres (article 51 RGPD) i les multes administratives i sancions, previstes en els articles 83 i 84 de l'RGPD.

⁹⁶⁵ KAMINSKI, M.E., "Binary Governance...", *op. cit.*, 2019, pàg. 1576.

i beneficis generals o miri pels interessos de les poblacions més vulnerables⁹⁶⁶. Tanmateix, un dels principals arguments que s'esgrimeix en contra la intervenció estatal és el risc que imposar regulacions molt restrictives en l'ús d'algoritmes pot ocasionar contratemps en el desenvolupament de la tecnologia. Si els costos associats al desenvolupament de la tecnologia esdevenen molt alts atesos els requisits normatius, els incentius per invertir en la tecnologia desapareixeran i es perdran els beneficis socials i econòmics del desenvolupament dels sistemes automatitzats⁹⁶⁷.

Un sistema de governança col·laborativa ha de tenir un sistema de control i equilibri⁹⁶⁸, actualment inexistent en l'anàlisi massiva de dades. En molts casos hi ha molt poca o directament no hi ha regulació de les interaccions entre la persona que dissenya l'algoritme, qui supervisa com s'executa i qui jutja si és correcte o no (equivalent a la separació de poders). Per tant, no hi ha sistema de control i equilibri que assegurí que no hi ha biaixos en el sistema, especialment crucial en un sistema d'execució⁹⁶⁹. Hi ha un desequilibri entre el poder del sector públic i del sector privat⁹⁷⁰.

Com ja hem esmentat, l'RGPD tal com està implementat actualment no sembla establir un sistema de rendició de comptes adequat de cara al públic general i inclús de cara a

⁹⁶⁶ SOLOW-NIEDERMAN, A., "Administering Artificial Intelligence", *op. cit.*, 2020, pàg. 685. L'autora també planteja la possibilitat d'un jurament hipocràtic -com el mèdic- en l'àmbit del desenvolupament professional al voltant de les dades en intel·ligència artificial que podria fomentar una identitat compartida, aclarir unes expectatives culturals mínimes i imbuir-los una consciència moral.

⁹⁶⁷ SORIANO ARNANZ, A., *Data Protection for the prevention...*, *op. cit.*, 2021, pàg. 192.

⁹⁶⁸ En anglès "checks and balances".

⁹⁶⁹ CRAWFORD, K. i SCHULTZ, J., "Big data and due process...", *op. cit.*, 2014, pàg. 120. En aquest article, Crawford i Schultz plantegen un "procés degut de les dades" que pretén, en comptes d'intentar regular la recollida de les dades, l'ús o divulgació *ex ante*, regular l'equitat dels processos analítics de les dades en relació amb com s'usen les dades personals en qualsevol procés, incloent processos en els quals l'anàlisi massiva de dades s'empra per determinar atributs o categories per una persona. Serien un conjunt de drets al procés degut que podrien exercir-se en el moment en què una companyia asseguradora denegés cobertura en base a l'anàlisi massiva de dades o si una persona qui contracta emprés l'anàlisi de dades per seleccionar a una la persona candidata (pàg. 109). Troben la problemàtica que aquest dret no pugui ser acomplert per moltes persones usuàries de l'anàlisi massiva de dades que no saben que ho són en el cas de les oportunitats per les quals no van ser seleccionades, la publicitat que no els va aparèixer en els resultats de la cerca i les feines que mai van saber que existien ja que no encaixaven en el perfil. Consideren, llavors, que es necessita una perspectiva del procés degut que inclogui supervisió i auditoria dirigida per part d'agències públiques (pàg. 124).

⁹⁷⁰ ZUIDERVEEN, F., *Discrimination, artificial intelligence...*, *op. cit.*, 2018, pàg. 44. Aquest autor enumera les carències de l'RGPD com a eina enfront la discriminació algorítmica derivades, entre altres causes, pel seu dèficit de compliment i execució, atès que les autoritats de protecció de dades tenen recursos i poders limitats per imposar sancions; SCHREURS, W., "Cogitas, Ergo Sum. The role of data protection...", *op. cit.*, 2008, pàg. 268. En aquest article es va més enllà i s'exposa que l'aplicació de la legislació de protecció de dades és una àrea notòriament difícil on les infraccions són rarament observades i perseguides. La solució tècnica proposada no és convincent per donar resposta a un problema essencialment sociocultural.

persones expertes. Hi ha una manca de transparència⁹⁷¹ combinada amb la manca de mecanismes o implicació de terceres parts que amenaça tant substantivament com processalment la legitimitat de l'RGPD com un règim de governança col·laborativa. Cal apuntar, igualment, que la governança col·laborativa presenta mètodes per incloure la comunitat però sovint són insatisfactoris. La necessitat de la implicació de la comunitat ve acompanyada de l'absència d'assistència i finançament per participar de manera significativa, la manca de mètodes per decidir quins grups rellevants són i qui els representa, la resistència de les entitats privades a incloure les parts interessades i la lentitud i complexitat del procés⁹⁷².

L'RGPD sovint dirigeix a les empreses la tasca de generar el contingut sobre com els drets individuals seran implementats i com abordar la manca d'equitat i les preocupacions relatives a les discriminacions algorítmiques. Però, en absència de supervisió pública, com podem assegurar que la defensa dels drets individuals i la governança col·laborativa estan treballant per l'interès general?⁹⁷³ Hi ha formes de govern i supervisió de terceres parts, com s'estableix en l'articulat, considerants i guies de l'RGPD, adequats per assegurar una governança col·laborativa de qualitat en absència de veritable transparència pública?⁹⁷⁴ A més, s'ha de tenir en compte que la relació entre “normes i espai” és cada cop més rellevant en la societat algorítmica. Els actors no estatals, corporacions privades i institucions de governança supranacional contribueixen a definir les normes i els codis de conducta, l'abast dels quals se solapa amb l'expressió tradicional de poder nacional sobirà⁹⁷⁵ i la supervisió que aquest poder, delimitat territorialment, pot proporcionar.

En definitiva, la governança col·laborativa subjacent en l'RGPD pot ser un bon sistema per abordar les discriminacions algorítmiques, ja que inclou els recursos i l'expertesa del sector privat en la regulació pública -elements essencials per abordar les noves

⁹⁷¹ KAMINSKI, M.E., “Binary Governance...”, *op. cit.*, 2019, pàg. 1607 i 1608. L'autora exemplifica la manca de transparència en els següents casos: el procés que instaura per establir codis de conducta no requereix transparència o implicació de terceres parts; malgrat que s'estableix la necessitat de terceres parts per ajudar a supervisar el compliment, aquestes no tenen poders de recollir informació sobre les companyies; el sistema de certificació, manca de rendició de comptes pública i; les avaluacions d'impacte (que només s'han de dur a terme en casos de risc alt), únicament involucren terceres parts en la fase de creació.

⁹⁷² SELBST, A. D., “An institutional view of Algorithmic Impact Assessments”, *op. cit.*, 2021, pàg. 160 i 161.

⁹⁷³ KAMINSKI, M. E. i MALGIERI, G., “Algorithmic impact assessments...”, *op. cit.*, 2021, pàg. 133.

⁹⁷⁴ KAMINSKI, M.E., “Binary Governance...”, *op. cit.*, 2019, pàg. 1609.

⁹⁷⁵ DE GREGORIO, G., *Digital Constitutionalism in Europe. Reframing Rights and Powers in the Algorithmic Society*, Cambridge, Cambridge University Press, 2022, pàg. 311.

tecnologies- que el sector públic no té. Endemés, la interpretació dels drets individuals com a elements per a incentivar les organitzacions a disminuir els riscos de discriminació dels sistemes algorítmics que empren, a través d'una responsabilitat proactiva amb un impacte indefectiblement sistèmic, pot proporcionar una combinació molt positiva per abordar la discriminació algorítmica. No obstant això, cal ser cautelosos amb la diferència entre el concepte teòric i la seva aplicació pràctica, ja que hi ha riscos molt rellevants i crucials per a la protecció de drets fonamentals: que el sector privat es converteixi *de facto* en autoregulator, ja sigui perquè amb la seva expertesa i recursos pot definir i incidir discrecionalment en conceptes clau de la regulació, o ja sigui perquè el poder públic no pot instaurar una rendició de comptes eficient -en haver-hi un desequilibri clar de poder entre sector públic i sector privat, i en no incloure en l'equació elements de transparència pública que facin participar a terceres parts, com la societat civil. També, els drets individuals han de ser efectius, atès que de no ser així, perden completament el seu factor dissuasiu de conductes discriminatòries.

A continuació estudiarem l'exigència de qualitat de les dades com a sistema per evitar que es generin certs biaixos algorítmics.

2. Qualitat de les dades per abordar els biaixos algorítmics

Com hem dit en el segon apartat del capítol III, en relació amb els biaixos algorítmics, l'RGPD no inclou aquesta problemàtica en l'articulat (s'hi fa referència no explícita en el considerant 71 de la norma que planteja la implementació de mesures tècniques adequades per resoldre errors i inexactituds en les dades, sense especificar quines) i la perspectiva que s'instaura específicament en l'article 9 de la norma estableix una protecció en base a dades sensibles que pot dificultar la protecció de la discriminació algorítmica, atès que totes les dades implicades en el tractament amb sistemes algorítmics poden esdevenir sensibles. Per consegüent, en aquest apartat analitzarem les eines de l'RGPD per assegurar la qualitat de les dades com a instruments per evitar (fins al punt que sigui possible) els biaixos algorítmics (és evident que excloem d'aquesta llista aquells que són implícits i dels quals la persona que els crea no en pot ser conscient). Per tant, cal tenir en compte que l'ús de mesures tècniques pot mitigar però no impedir l'existència de

biaixos algorítmics i que, a més, els processos i mètriques per a un aprenentatge automàtic just, responsable i transparent no sempre es tradueixen fàcilment als marcs legals⁹⁷⁶.

Els elements que proposa l'RGPD són l'exigència d'exactitud de les dades (2.1), i els drets que atorga a les persones per controlar i actualitzar les seves dades (2.2). També analitzem la proposta que s'ha fet per part de la doctrina sobre capitalitzar les dades per poder tenir un major control sobre elles (2.3).

2.1. L'exactitud⁹⁷⁷

L'exactitud de les dades és un principi que s'estableix en l'article 5.4 de l'RGPD i que precisa que s'adoptaran totes les mesures raonables per a què se suprimeixin -dret a l'oblit⁹⁷⁸-, o es rectifiquin (correguin o actualitzin) totes les dades inexactes respecte a les finalitats per les quals són tractades. Aquesta exactitud es requereix en totes les etapes en les quals estiguin involucrades les dades personals.

En el cas de tractaments de dades personals que comporten sistemes d'intel·ligència artificial, l'existència de biaixos en els models està lligada a l'exactitud o la qualitat de les dades⁹⁷⁹. Les dades associades a les persones interessades que o bé siguin directament recollides, o bé siguin deduïdes, han de ser exactes. La persona responsable de tractament ha d'utilitzar procediments matemàtics o estadístics adequats per l'elaboració de perfils garantint que les dades que s'utilitzen són exactes⁹⁸⁰. També es requereix que els procediments emprats per a la inferència d'informació sobre una persona interessada siguin precisos⁹⁸¹, de manera que el procés que es dugui a terme sigui predictable i estable.

⁹⁷⁶ COBBE, J. "Administrative law and the machines of government: judicial review of automated public-sector decision making", *Legal Studies*, vol. 39, núm. 4, 2019, pàg. 638.

⁹⁷⁷ S'entén per "exactitud" segons l'article 5.1.d de l'RGPD, que les dades personals siguin exactes, i si fos necessari actualitzades; que s'adoptin totes les mesures raonables per tal que es suprimeixin o rectifiquin sense dilació les dades personals que siguin inexactes respecte les finalitats per les quals són tractades.

⁹⁷⁸ Amb anterioritat a l'entrada en vigor de l'RGPD, aquest dret va ser reconegut pel TJUE en relació als motors de recerca d'internet. Vegeu: TJUE (Gran Sala), sentència de 13 de maig de 2014, C-131/12, Google Spain S.L. contra Agència Espanyola de Protecció de dades, par. 82.

⁹⁷⁹ Es precisa en l'article 47.2.d de l'RGPD que l'autoritat de control competent aprovarà normes corporatives que hauran d'especificar com a mínim certs elements, entre ells, la qualitat de les dades.

⁹⁸⁰ Considerant 71 de l'RGPD.

⁹⁸¹ AEPD, *Adecuación al RGPD de tratamientos que incorporan Inteligencia Artificial. Una introducción*, Madrid, 2020, pàg. 36.

Hi ha diversos factors que poden influir en l'exactitud, com per exemple la pròpia implementació dels sistemes d'intel·ligència artificial (que elements externs introdueixin errors en les dades d'entrada o que existeixin errors de programació o de disseny que traslladin de forma errònia el model a la seva implementació pràctica); que el conjunt de dades d'entrenament o validació estigui viciat amb errors, informació deliberadament errònia o biaixos que impossibilitin que les inferències siguin correctes; o que, quan ens referim a sistemes d'intel·ligència artificial que utilitzin tècniques adaptatives, aquests estiguin essent utilitzats principalment per un grup determinat de subjectes amb característiques particulars que introdueixin nous biaixos de realimentació⁹⁸².

Quan ens referim a les dades d'entrenament, s'han de dur a terme tècniques de depuració i traçabilitat per assegurar la fidelitat i integritat del conjunt de les dades. En alguns models es poden dividir les dades d'entrenament entre dades “dures” (objectives, generalment quantificables -qualificacions, percentatge d'assistència, valors analítics, resultats de proves, etc.-) i dades “toves” (subjectives, generalment qualitatives).⁹⁸³

Kullman, atesa la dificultat d'aportar indicis de discriminació, planteja la possibilitat que una persona pugui presentar prova a través d'informació obtinguda per les dades d'entrenament o de prova del sistema⁹⁸⁴. També s'ha plantejat la idea que les persones que creïn una base de dades han d'explicar les seves característiques i les seves limitacions, per exemple, el propòsit i el context pel qual es van crear, quin pretractament o rentat de dades es va fer, o la demografia representada en les dades. Aquesta informació hauria de permetre escollir a les persones que volen emprar-les, les bases de dades de manera eficient, incrementant la qualitat del sistema d'intel·ligència artificial. Si les bases de dades estan etiquetades i algú compra un algoritme d'ajuda a la contractació, poden potencialment escollir entre models entrenats en diferents indústries o en grups de persones amb diferents demografies que s'adaptin el millor possible al grup de persones candidates⁹⁸⁵, i per consegüent, aportin uns resultats menys discriminatoris.

⁹⁸² *Ídem, ibidem.*

⁹⁸³ *Ídem, ibidem.*

⁹⁸⁴ KULLMANN, M., “Platform Work, Algorithmic Decision Making, and EU Gender Equality Law”, *International Journal of Comparative Labour Law and Industrial Relations*, vol. 34, núm. 1, 2018, pàg. 15.

⁹⁸⁵ SELBST, A. D., “An institutional view of Algorithmic Impact Assessments”, *op. cit.*, 2021, pàg. 160 i 186. Versions d'aquest etiquetatge de dades es van proposar per part de tres equips diferents, i cadascun va emprar una metàfora diferent per la documentació però les conclusions van ser similars. Vegeu: GEBRU, T., MORGENSTERN, M., VECCHIONE, B., WORTMAN VAUGHAN, J., WALLACH, H., DAUMÉ III, H. i CRAWFORD, K., “Datasheets for datasets”, *Communications of the ACM*, vol. 64, núm. 12, 2021,

Quant al control de les persones sobre les seves dades personals que poden contribuir a l'exactitud de les dades, l'RGPD confereix diversos drets.

2.2. Drets de les persones sobre les seves dades: accés, supressió, rectificació, limitació del termini de conservació i portabilitat

Els drets que confereix l'RGPD a les persones sobre les seves dades són el dret d'accés, supressió, rectificació, limitació del termini de conservació i portabilitat (també farem una menció al bloqueig de dades només establert en la LOPDGDD).

La doctrina es refereix als drets individuals que confereix l'RGPD que acabem de referenciar com a aspectes necessaris de l'oportunitat d'un individu a ser escoltat. Aquests drets permeten a les persones de participar en la formació de la seva "objectificació" incloent la correcció i supressió de la informació. Idealment, permetent aquesta participació, aquests drets potencialment creen un diàleg entre una persona a la qual pertanyen les dades i l'organització que les utilitza, sobre la racionalitat de les decisions algorítmiques i altres formes d'elaboració de perfils i tractament de dades, servint bàsicament per corregir errors factuais o d'inferència⁹⁸⁶. No obstant això, també sembla rellevant puntualitzar que poden presentar unes premisses d'exercici efectiu molt estàtiques pels sistemes algorítmics, on les dades flueixen d'una manera dinàmica i ràpida. També, com esmenten Edwards i Veale, evidencien com es depèn en excés de drets individuals pel que molt sovint són danys col·lectius⁹⁸⁷. A més, s'ha criticat que la perspectiva d'autodeterminació informativa (basar el tractament en el consentiment de les persones interessades i establir un catàleg de drets que les persones sotmeses al tractament poden exercir per protegir els seus interessos) no ha generat un sistema efectiu

pàg. 86-92; HOLLAND, S., HOSNY, A., NEWMAN, S., JOSEPH, J. i CHMIELINSKI, K., "The Dataset Nutrition Label: A Framework To Drive Higher Data Quality Standards", 2018, pàg. 1-21. Disponible a: <https://arxiv.org/>. Darrera visita el 2 de setembre de 2023; BENDER, E. M., i FRIEDMAN, B., "Data Statements for Natural Language Processing: Toward Mitigating System Bias and Enabling Better Science", *Transactions of the Association for Computational Linguistics*, vol. 6, 2018, pàg. 587-604.

⁹⁸⁶ KAMINSKI, M.E., "Binary Governance...", *op. cit.*, 2019, pàg. 1592.

⁹⁸⁷ VEALE, M. i EDWARDS, L., "Slave to the Algorithm? Why...", *op. cit.*, 2017, pàg. 81.

de control⁹⁸⁸. Creiem pertinent, tanmateix, descriure'ls, ja que actualment són el règim que dona més control a les persones sobre les seves dades, malgrat les seves limitacions.

En primer lloc, en relació amb el dret d'accés (article 15 RGPD), la persona interessada podrà consultar les dades que sobre ella té la persona responsable de tractament⁹⁸⁹. Malgrat que la norma general es refereix al conjunt de facultats derivades del dret a la protecció de dades, el dret d'accés a les dades personals és un dret de configuració independent, és a dir, que podrà exercir-se de forma independent als altres drets. És la persona titular de les dades qui decideix en últim terme si l'exerceix de forma conjunta i independent als altres drets. Malgrat això, no es pot negar que en ocasions és necessari en primer lloc el dret d'accés per poder exercir la resta de drets⁹⁹⁰: només si conec les dades que tracta sobre mi la persona responsable, podré, per exemple, oposar-me, rectificar-les o sol·licitar la seva supressió. A més, és la persona titular del dret qui decideix també el mitjà de realitzar la seva petició d'accés, condicionant així, la forma amb què se li donarà resposta⁹⁹¹.

La tutela del dret d'accés no té cap peculiaritat respecte a la resta de facultats garantides pel dret de protecció de dades personals. Per exemple, davant d'una denegació total o parcial de la sol·licitud, o en absència d'una resposta per part de la persona responsable, la persona interessada haurà de sol·licitar la tutela a l'autoritat competent, amb el procediment previst en el capítol VIII de l'RGPD (art. 77-84), que té el seu procediment homòleg a escala nacional en el Títol VIII LOPDGDD (art. 63-69) i Capítol I i II del Títol IX (art. 115-116 i 117-119) del Real Decreto 1720/2007, de 21 de diciembre, por el que se aprueba el Reglamento de desarrollo de la Ley Orgánica 15/1999, de 13 de diciembre,

⁹⁸⁸ SORIANO ARNANZ, A., "La propuesta de Reglamento..." *op. cit.*, 2021, pàg. 6.

⁹⁸⁹ ARENAS RAMIRO, M., "El derecho de acceso y las condiciones generales de ejercicio de los derechos (Comentario al artículo 15 RGPD y a los artículos 12 y 13 LOPDGDD)", a TRONCOSO REIGADA, A., (dir.) i VILLALBA CANO, L., (coord.), *Comentario al Reglamento General de Protección de Datos y a la Ley Orgánica de Protección de Datos Personales y Garantía de los Derechos Digitales*, Tomo I, Pamplona, Aranzadi, 2021, pàg. 1470.

⁹⁹⁰ CITRON, D. K. i PASQUALE, F., "The Scored Society: Due Process for Automated Predictions", *Washington Law Review*, vol. 89, núm. 1, 2014, pàg. 16. El dret d'accés es considera un pilar pel control per part de les persones consumidores de les seves puntuacions de crèdit. En l'article referenciat es discuteix l'abast de la Fair Credit Reporting Act (FCRA), que és la norma del govern federal dels Estats Units d'Amèrica per promoure la precisió, la imparcialitat i privacitat de la informació de la persona consumidora continguda en els arxius de les agències d'informes sobre ella. Expliquen que es va aprovar la llei per la preocupació creixent relativa a les maneres d'usar de manera invisible i perjudicial per la persona consumidora les bases de dades personals. Aquesta llei obliga a les oficines de crèdit a assegurar-se que els seus historials són precisos i dona a les persones consumidores, entre altres, el dret a inspeccionar els seus historials i a demanar correccions.

⁹⁹¹ ARENAS RAMIRO, M., "El derecho de acceso y las condiciones...", *op. cit.*, 2021, pàg. 1465.

de protecció de dades de caràcter personal⁹⁹² (d'ara en endavant, RLOPD), vigent en allò que no contradigui l'RGPD i la LOPDGDD⁹⁹³.

En segon lloc, el dret a la rectificació de les dades, regulat a l'article 16 de l'RGPD, disposa que la persona interessada podrà obtenir de la persona responsable del tractament, sense dilació indeguda, una rectificació de les dades personals que li concerneixen⁹⁹⁴. El dret es concreta en què es modifiqui el contingut material de les dades en el sentit sol·licitat per la persona interessada (art. 31.1 RLOPD)⁹⁹⁵. Cal puntualitzar que en els sistemes algorítmics la rectificació de dades ja emprades en el tractament no tindrà efecte retroactiu, i es podran seguir emprant de manera no acurada.

En tercer lloc, quant al dret de supressió de dades personals, el redactat de l'article 17 de l'RGPD estipula que es tracta d'un dret l'exercici del qual es pot fer efectiu sota certes condicions. Segons l'article 17, o bé les dades personals ja no són necessàries en relació amb les finalitats per a les quals van ser recollides o són tractades d'una altra manera; o bé la persona interessada retira el consentiment sota el qual es basa el tractament i aquest no es basa en un altre fonament jurídic; o bé la persona interessada s'oposa al tractament i no prevalen altres motius legítims per al tractament; o bé les dades personals han estat tractades de manera il·lícita; o bé les dades personals s'han de suprimir pel compliment d'una obligació legal establerta pel dret de la UE o dels estats membres que s'aplica al responsable de tractament; o bé, finalment, si les dades personals s'han obtingut en relació amb l'oferta de serveis de la societat de la informació⁹⁹⁶.

No només la persona interessada s'ha d'adscriure a alguna condició de les aquí enumerades, sinó que no es podria invocar el dret quan les dades personals siguin necessàries per exercir el dret a la llibertat d'expressió i informació; pel compliment d'una

⁹⁹² Real Decreto 1720/2007, de 21 de diciembre, por el que se aprueba el Reglamento de desarrollo de la Ley Orgánica 15/1999, de 13 de diciembre, de protección de datos de carácter personal. BOE, núm. 17, de 19 de enero de 2008. España.

⁹⁹³ ARENAS RAMIRO, M., "El derecho de acceso y las condiciones...", *op. cit.*, 2021, pàg. 1490.

⁹⁹⁴ Article 16 de l'RGPD.

⁹⁹⁵ PASQUAL HUERTA, P., "El derecho de rectificación (Comentario al artículo 16 RGPD)", a TRONCOSO REIGADA, A., (dir.) i VILLALBA CANO, L., (coord.), *Comentario al Reglamento General de Protección de Datos y a la Ley Orgánica de Protección de Datos Personales y Garantía de los Derechos Digitales*, Tomo I, Pamplona, Aranzadi, 2021, pàg. 1535.

⁹⁹⁶ Article 17.1 de l'RGPD. A la doctrina espanyola, vegeu: DÍAZ DÍAZ, E., "El derecho de supresión. El derecho al olvido (Comentario al artículo 17 RGPD y artículo 15 LOPDGDD)", a TRONCOSO REIGADA, A., (dir.) i VILLALBA CANO, L., (coord.), *Comentario al Reglamento General de Protección de Datos y a la Ley Orgánica de Protección de Datos Personales y Garantía de los Derechos Digitales*, Tomo I, Pamplona, Aranzadi, 2021, pàg. 1561-1604.

obligació legal que requereixi el tractament de dades imposat pel dret de la UE o dels estats membres que s'apliqui al responsable de tractament, o d'una missió realitzada en interès públic o en l'exercici de poders públics atorgats al responsable; per raons d'interès públic en l'àmbit de la salut pública, per finalitats d'arxiu en interès públic, d'investigació científica, històrica o estadística; o per la formulació, exercici o defensa de reclamacions⁹⁹⁷.

També, sembla discutible que no es tracti d'un dret que es pugui exercir sense haver de justificar el compliment d'una de les condicions enumerades. A més, crea molts dubtes l'adaptació que pot tenir quan ens referim a decisions automatitzades, o amb més precisió, a sistemes d'intel·ligència artificial d'aprenentatge automàtic on les dades s'han utilitzat per entrenar la màquina o formen part del sistema decisiu. Addicionalment, sembla un dret difícil d'assolir si es té en compte que la revocació del consentiment no es considera amb efectes retroactius, de manera que tot l'ús que se li hagi donat a les dades personals ja no es podrà esborrar, tal com ja hem esmentat en relació amb el dret a la rectificació.

S'afegeix que l'RGPD no recomana marcs tècnics específics o mètodes de privadesa per preservar els requisits legals que introdueix per fer-lo efectiu, de manera que és més difícil el seu compliment⁹⁹⁸. És pertinent mencionar que l'article 32 de la LOPDGDD⁹⁹⁹ planteja també el dret al bloqueig, no previst a l'RGPD -que es basa en la identificació i reserva de dades per impedir el seu tractament, excepte per posar-les a disposició de jutges/ses i tribunals, el Ministeri Fiscal o les administracions públiques competents, en particular les autoritats de protecció de dades, per a l'exigència de possibles responsabilitats derivades del tractament i només pel termini de prescripció de les mateixes.

En quart lloc, s'estableix també com a principi la limitació del termini de conservació de les dades (article 18 de l'RGPD), de forma que no es permeti la identificació de la persona interessada més temps del necessari per a la finalitat determinada i que les dades siguin tractades de manera que es garanteixi una seguretat i integritat adequada de les dades personals, inclosa la protecció contra el tractament no autoritzat o il·lícit i contra la pèrdua, destrucció o dany accidental mitjançant les mesures tècniques o organitzatives

⁹⁹⁷ Article 17.3 de l'RGPD.

⁹⁹⁸POLITOU, E. *et al.*, "Backups and the right to be forgotten in the GDPR: An uneasy relationship", *Computer Law & Security Review: The International Journal of Technology Law and Practice*, vol. 34, núm. 6, 2018, pàg. 1256.

⁹⁹⁹Article 32 de l'LOPDGDD.

apropiades¹⁰⁰⁰. Les autoritats de control poden reclamar a la persona responsable de tractament que atengui o contesti una sol·licitud de limitació del tractament¹⁰⁰¹ (també les mateixes sol·licituds sobre altres drets).

L'article 19 de l'RGPD disposa que la persona responsable de tractament comunicarà qualsevol rectificació o supressió de dades personals o limitació del tractament efectuat en relació a l'article 16, l'article 17.1 i l'article 18, a cada un dels destinataris als quals se'ls hi hagin comunicat les seves dades personals, excepte que sigui impossible o exigeixi un esforç desproporcionat¹⁰⁰². Aquesta última frase deixa marge d'actuació a la persona responsable de tractament, que, si no duu a terme la seva obligació, haurà de justificar perquè ha estat impossible o ha exigut un esforç desproporcionat.

En cinquè lloc, l'article 20 de l'RGPD precisa el dret a la portabilitat, que consisteix en rebre les dades personals que incumbeixin a la persona interessada, que aquesta hagi facilitat a una persona responsable de tractament, en format estructurat, d'ús comú i lectura mecànica, i transmetre-les a una altra persona responsable de tractament sense que la primera responsable de tractament pugui impedir-ho quan a) el tractament estigui basat en el consentiment o b) el tractament s'efectuï per mitjans automatitzats¹⁰⁰³.

En sisè lloc, finalment, el dret de l'article 21 de l'RGPD a oposar-se en qualsevol moment al tractament de dades el desenvolupen l'article 34.c i 36 del RLOPD¹⁰⁰⁴. L'oposició no tindrà lloc quan el tractament sigui necessari per l'execució d'un contracte, estigui autoritzat pel dret de la UE o s'hagi atorgat el consentiment explícit. Les guies del Grup de Treball de l'Article 29, suggereixen que el dret a oposar-se ha de ser un procés intern de la companyia (no hi ha indicacions sobre què pot implicar aquest procés). Això planteja la qüestió de si una companyia els interessos de la qual no sempre s'alineen amb els seus usuaris és capaç de proveir-se d'un procés adequat i resultats justos, fet que suposa la

¹⁰⁰⁰ Article 5.4 de l'RGPD.

¹⁰⁰¹ PASQUAL HUERTA, P., "El derecho a la limitación de tratamiento (Comentario al artículo 18 RGPD)", a TRONCOSO REIGADA, A., (dir.) i VILLALBA CANO, L., (coord.), *Comentario al Reglamento General de Protección de Datos y a la Ley Orgánica de Protección de Datos Personales y Garantía de los Derechos Digitales*, Tomo I, Pamplona, Aranzadi, 2021, pàg. 1631.

¹⁰⁰² Article 19 de l'RGPD.

¹⁰⁰³ Article 20 de l'RGPD. Per aprofundir en el dret a la portabilitat, vegeu: SANTAMARÍA RAMOS, F.J., "El derecho a la portabilidad: Un nuevo derecho esencial en materia de protección de datos (Comentario al artículo 20 RGPD y al artículo 17 LOPDGD)", a TRONCOSO REIGADA, A., (dir.) i VILLALBA CANO, L., (coord.), *Comentario al Reglamento General de Protección de Datos y a la Ley Orgánica de Protección de Datos Personales y Garantía de los Derechos Digitales*, Tomo I, Pamplona, Aranzadi, 2021, pàg. 1655-1680.

¹⁰⁰⁴ Article 34.c i 36 de l'RLOPD.

necessitat de desenvolupar polítiques per donar cos a aquest dret¹⁰⁰⁵. Si la persona interessada finalment no obté resposta expressa de la companyia, podrà interposar una reclamació a l'autoritat de control¹⁰⁰⁶.

Malgrat que, com hem exposat al llarg de l'apartat, els drets que confereix l'RGPD a les persones per controlar les seves dades tenen certs conflictes per adaptar-se de manera precisa quan es veuen involucrats sistemes algorítmics, generen eines de control o si més no de comunicació entre les persones que veuen tractades les seves dades i qui realitza el tractament, de manera que poden participar a garantir uns mínims de qualitat de les dades.

Analitzem, a continuació, la proposta que s'ha fet per part de la doctrina per capitalitzar les dades com a forma de millorar el control que la persona pot tenir d'elles i la incidència d'aquesta proposta en la lluita envers la discriminació.

2.3. Capitalitzar les dades

S'ha plantejat com una alternativa legal, amb la finalitat de potenciar el control individual sobre les pròpies dades, la idea de propietat de les dades personals. Donat que en l'economia digital la informació sobre els individus es considera que té sovint, i cada cop més per a qui participa en el mercat, un valor comparable als diners, s'ha plantejat la mercantilització de les dades, és a dir, una vessant patrimonial de la protecció de dades, creant la possibilitat que aquella persona que cedeixi les dades n'obtingui una remuneració.

És innegable el valor creixent que estan adquirint les dades i el seu ús com a moneda de canvi en molts negocis digitals que aparentment es presten de forma "gratuïta" a les persones consumidores¹⁰⁰⁷. Les empreses que posseeixen les dades s'han atorgat *de facto*

¹⁰⁰⁵ ARTICLE 29 DATA PROTECTION WORKING PARTY, *Guidelines on Automated individual...*, *op. cit.*, 2018, pàg. 18.

¹⁰⁰⁶ ARENAS RAMIRO, M., "El derecho de oposición (Comentario al artículo 21 RGPD y al artículo 18 LOPDGDD)", a TRONCOSO REIGADA, A., (dir.) i VILLALBA CANO, L., (coord.), *Comentario al Reglamento General de Protección de Datos y a la Ley Orgánica de Protección de Datos Personales y Garantía de los Derechos Digitales*, Tomo I, Pamplona, Aranzadi, 2021, pàg. 1723.

¹⁰⁰⁷ PERETE RAMÍREZ, C. i GARCÍA MEXÍA, P., "La propiedad sobre el dato. ¿Cabe una vertiente patrimonial de la protección de datos?", a TRONCOSO REIGADA, A., (dir.) i VILLALBA CANO, L.,

un dret sobre aquestes, que les persones usuàries han acceptat pacíficament, ja que entenien que era el preu que havien de pagar per rebre serveis gratuïts¹⁰⁰⁸.

Malgrat l'RGPD ha atorgat a les persones un major control sobre les seves dades, no ha arribat al punt d'atribuir-ne un dret a la seva propietat¹⁰⁰⁹. Nombroses persones del món acadèmic han defensat que el dret a la portabilitat de dades és el que més s'adapta a l'enfocament del dret a la propietat de les dades. Segons el Grup de Treball de l'Article 29, l'objectiu principal de la portabilitat de dades és promoure el control individual i assegurar-se que es juga un rol actiu en l'ecosistema de dades¹⁰¹⁰. La limitació més important en l'àmbit d'aplicació d'aquest dret seria la prohibició de renunciar a les garanties de la protecció de dades¹⁰¹¹.

Tot i les possibilitats que obre una vessant patrimonial de les dades en relació amb el control que aquesta interpretació pot proporcionar a aquelles persones que cedeixen les seves dades, les dades personals no són fàcils de concretar en els drets de propietat individual. Precisament, no es pot analitzar des de la perspectiva que concerneix només a un individu, sinó que inevitablement tindrà efectes en d'altres, per tant, el control individual sobre les dades personals i l'exercici efectiu dels drets de propietat individual són difícils d'aconseguir¹⁰¹². En referència a la lluita envers la discriminació, un control de les dades basat en una vessant patrimonial de les dades personals no sembla oferir en l'actualitat una millor gestió que aquella establerta a l'RGPD.

Resumidament, tant l'exigència d'exactitud com els articles 15, 16, 17, 18, 20 i 21 de l'RGPD -que confereixen drets a la persona que veu tractades les seves dades perquè pugui tenir-ne un control més exhaustiu-, poden participar a millorar la qualitat de les dades que tenen una incidència directa en els biaixos algorítmics causants de la discriminació algorítmica. No obstant això, cal apuntar que l'RGPD planteja la qualitat de dades de manera estàtica, és a dir, les dades que hi ha es poden canviar, fet que no s'adapta als sistemes algorítmics que evolucionen a través de la correlació de dades que no es podran alterar retrospectivament. A més, només genera eines per fer front a la

(coord.), *Comentario al Reglamento General de Protección de Datos y a la Ley Orgánica de Protección de Datos Personales y Garantía de los Derechos Digitales*, Tomo I, Pamplona, Aranzadi, 2021, pàg. 1700.

¹⁰⁰⁸ *Ídem*, pàg. 1685.

¹⁰⁰⁹ *Ídem*, pàg. 1700.

¹⁰¹⁰ PURTOVA, N., "Do property rights in personal data...", *op. cit.*, 2017, pàg. 70.

¹⁰¹¹ *Ídem*, pàg. 68-69.

¹⁰¹² *Ídem*, pàg. 77.

vessant tècnica d'aquesta problemàtica, i no podem ignorar la vessant social i psicològica, que fa necessàries diferents accions per mitigar l'aparició de biaixos i reduir-ne l'impacte, com la sensibilització social envers els biaixos inconscients, equips de treball amb diversitat d'edat, ètnia, cultura o gènere, ús de dades el més representatives possible, promoure IA diversa i inclusiva, comprendre i mesurar biaixos o generar noves bases de dades d'entrenament¹⁰¹³.

Tot seguit estudiem de quina manera l'accés a informació i les exigències de transparència tal com estan plantejades a l'RGPD poden participar a millorar la protecció envers la discriminació algorítmica.

3. Informació útil i senzilla per fer front a la discriminació algorítmica

En aquest apartat analitzem sota quines condicions els requisits de transparència de l'RGPD -sobre els quals hem abordat els seus límits-, poden ser útils per a les persones que han patit discriminació.

Com hem dit en l'apartat 3 del capítol III, els obstacles principals per a la utilitat -en relació amb la discriminació algorítmica- de la informació que es dona a la persona que veu tractades les seves dades són: que no és clar quin contingut se li pot donar a la informació perquè sigui entenedora per a una persona que s'ha vist discriminada, i que pugui, en conseqüència, utilitzar-la per desafiar el sistema algorítmic; i que s'aborda des d'una perspectiva individual.

No hi ha cap altre norma que obligui a donar certa informació sobre la presa de decisions automatitzades, així que l'RGPD, si més no, dona l'oportunitat de reclamar. Plantegem, en primer lloc, el debat sobre el dret a l'explicació i quina seria la seva implicació en la lluita envers la discriminació (3.1); en segon lloc, de quina manera creiem que és millor proporcionar informació individual perquè la persona pugui fer-la servir en un context de

¹⁰¹³OBSERVATORI D'ÈTICA EN INTEL·LIGÈNCIA ARTIFICIAL DE CATALUNYA (OEIAC), *Intel·ligència artificial, ètica i societat...*, op. cit., 2021, pàg. 36. En relació amb la intervenció sobre les bases de dades, vegeu: GERARDS, J. i XENIDIS, R., *Algorithmic Discrimination in Europe...*, op. cit., 2021, pàg. 149-150. Les autores fan referència a la tècnica d'augment de dades (una estratègia per reforçar la diversitat de les dades en una base de dades donada) o estratègies de neteja, per ajudar a desesbiaixar les dades en el moment de selecció i preparació de les bases de dades.

discriminació (3.2); i finalment, en tercer lloc, analitzem com la idea d'una política de transparència que es dona en l'RGPD pot assentar la base per una bona regulació de la discriminació algorítmica (3.3).

3.1. El dret a una explicació en la protecció envers la discriminació

Certa part de la doctrina entén que en base al considerant 71¹⁰¹⁴, combinat amb l'anàlisi sistemàtica de l'article 13, 14 i 15 de l'RGPD¹⁰¹⁵, les persones que són subjectes d'una decisió únicament automatitzada tenen dret a una explicació de la decisió a la qual ha arribat el sistema¹⁰¹⁶. S'ha considerat que aquesta explicació no seria tant de la presa de decisió concreta per a una persona individual (subject-based explanation) sinó que hauria de reduir-se al funcionament general del model (model-based explanation)¹⁰¹⁷. Cal puntualitzar, en relació amb l'accés a la informació, que el Reglamento de Mercados Digitales¹⁰¹⁸, en el seu article 21.3, possibilita expressament a la Comissió Europea (en l'exercici de les seves funcions de regulació del mercat), per mitjà d'una sol·licitud o decisió (que ha d'indicar la finalitat i el termini d'entrega), exigir l'accés a qualsevol dada o algoritme de les empreses i informació sobre les proves, així com sol·licitar explicacions sobre aquesta informació¹⁰¹⁹.

¹⁰¹⁴Estableix que les persones interessades haurien de tenir el dret d'obtenir una explicació de qualsevol decisió que es produeixi després de l'avaluació de les seves dades personals.

¹⁰¹⁵Els articles 13.2.f i 14.2.g de l'RGPD estableixen que els responsables de tractament han de proveir a les persones interessades d'informació significativa sobre la lògica del sistema de decisió automatitzada, sobre la importància de l'automatització i les conseqüències previstes de la decisió, i l'article 15.1.e de l'RGPD autoritza a les persones interessades a obtenir aquest tipus d'informació de la persona responsable de dades.

¹⁰¹⁶WOLTERS, P.T.J., "The Control by and Rights of the Data Subject...", *op. cit.*, 2018, pàg. 12.

¹⁰¹⁷WACHTER, S., MITTELSTADT, B. i FLORIDI, L., "Why a Right to Explanation of Automated...", *op. cit.*, 2017, pàg. 76-99.

¹⁰¹⁸Reglamento (UE) 2022/1925 del Parlamento Europeo y del Consejo de 14 de septiembre de 2022 sobre mercados disputables y equitativos en el sector digital y por el que se modifican las Directivas (UE) 2019/1937 y (UE) 2020/1828 (Reglamento de Mercados Digitales). DOUE, L 265/1, 12.10.2022. Aquest Reglament té per objecte garantir que les grans plataformes en línia que actuen com a guardianes en els mercats digitals (per ser guardianes en els mercats digitals, el Reglament estableix un conjunt de criteris definits, que tenen en compte, per exemple, tenir un impacte significatiu en el mercat interior (art. 3.1.a) o una posició consolidada i duradora en el mercat (art. 3.1.c)) es comportin de manera equitativa en línia.

¹⁰¹⁹En el cas que empreses no facilitin l'accés a les bases de dades, algoritmes o informació sobre les proves en resposta a una sol·licitud, la Comissió Europea les podrà sancionar amb multes que no excedeixin l'1% del seu volum de negocis total a escala mundial en l'exercici anterior (art. 30.3.e), com també podrà imposar multes coercitives diàries que no excedeixin el 5% de la mitjana diària del volum de negocis a escala mundial en l'exercici anterior (art. 31.1.d), per obligar-les a donar accés.

Als Països Baixos, s'ha defensat que existeix una obligació general per les autoritats públiques d'assegurar l'explicabilitat, transparència i accessibilitat dels algorismes perquè les persones puguin entendre com s'han vist afectades i per donar-los les eines per impugnar de manera efectiva l'algoritme davant els tribunals¹⁰²⁰. A escala nacional, també es poden veure iniciatives per adaptar les legislacions a l'ecosistema digital, com per exemple, la Ley 12/2021, de 28 de septiembre, por la que se modifica el texto refundido de la Ley del Estatuto de los Trabajadores, aprobado por el Real Decreto Legislativo 2/2015, de 23 de octubre, para garantizar los derechos laborales de las personas dedicadas al reparto en el ámbito de plataformas digitales¹⁰²¹, que introdueix un apartat d) a l'article 64.4 de la Ley del Estatuto de los Trabajadores, en el qual s'especifica que el Comitè d'Empresa té dret a ser informat sobre els paràmetres, regles i instruccions en què es basen els algorismes i sistemes d'intel·ligència artificial que afecten la presa de decisions¹⁰²² o l'article XXV de la Carta de Drets Digitals (que no té caràcter normatiu), que exposa els drets davant la intel·ligència artificial, entre ells l'establiment de condicions d'explicabilitat¹⁰²³, i l'article XVIII on s'estableix la possibilitat que la llei reguli les condicions de transparència i accés al codi font, especialment amb l'objectiu de verificar que no produeix resultats discriminatoris¹⁰²⁴. En aquest sentit, el Consell Constitucional francès va plantejar, en la decisió de 12 de juny de 2018, la prohibició de certs tipus d'algorismes sobre els quals no es podia donar una explicació en relació al seu funcionament si ho requerien els usuaris. D'aquesta manera, es deixava de banda l'ús de

¹⁰²⁰ ABRvS, 17 May 2017, ECLI:NL:RVS:2017:1259, Aerius I; ABRvS, 18 July 2018, ECLI:NL:RVS:2018:2454, Aerius II.

¹⁰²¹ Ley 12/2021, de 28 de septiembre, por la que se modifica el texto refundido de la Ley del Estatuto de los Trabajadores, aprobado por el Real Decreto Legislativo 2/2015, de 23 de octubre, para garantizar los derechos laborales de las personas dedicadas al reparto en el ámbito de plataformas digitales. BOE, núm. 233, de 29 de septiembre de 2021. España.

¹⁰²² La lletra d) introduïda en l'article 64.4 de l'Estatut dels Treballadors, versa de la següent manera: «d) (El Comité de Empresa tiene derecho a) ser informado por la empresa de los parámetros, reglas e instrucciones en los que se basan los algoritmos o sistemas de inteligencia artificial que afectan a la toma de decisiones que pueden incidir en las condiciones de trabajo, el acceso y mantenimiento del empleo, incluida la elaboración de perfiles.»

¹⁰²³ La Carta de Drets Digitals espanyola estableix en el seu apartat XXV que “1. La inteligencia artificial deberá asegurar un enfoque centrado en la persona y su inalienable dignidad, perseguirá el bien común y asegurará cumplir con el principio de no maleficencia. 2. En el desarrollo y ciclo de vida de los sistemas de inteligencia artificial: a) Se deberá garantizar el derecho a la no discriminación cualquiera que fuera su origen, causa o naturaleza, en relación con las decisiones, uso de datos y procesos basados en inteligencia artificial. b) Se establecerán condiciones de transparencia, auditabilidad, explicabilidad, trazabilidad, supervisión humana y gobernanza. En todo caso, la información facilitada deberá ser accesible y comprensible. c) Deberán garantizarse la accesibilidad, usabilidad y fiabilidad. 3. Las personas tienen derecho a solicitar una supervisión e intervención humana y a impugnar las decisiones automatizadas tomadas por sistemas de inteligencia artificial que produzcan efectos en su esfera personal y patrimonial.”

¹⁰²⁴ Article XVIII.6.b de la Carta de Drets Digitals Espanyola.

mètodes d'aprenentatge automàtic amb l'habilitat de millorar el seu funcionament de manera autònoma i els algorismes protegits per drets de propietat intel·lectual¹⁰²⁵.

Goodman i Flaxman, adduint que de l'RGPD es desprèn un dret a l'explicació, defensen que qualsevol explicació adequada, com a mínim, rendirà compte sobre com les dades d'entrada es relacionen amb les prediccions, possibilitant respondre a qüestions com si el model és més o menys procliu a recomanar un préstec si la persona que el demana pertany a un grup oprimat, o quines dades tenen un pes més gran en la predicció d'un algoritme¹⁰²⁶. *A contrario sensu*, Castets-Renard entén que la informació que conscienciï la persona sobre el tipus de tractament que s'està duent a terme esdevé impossible a causa de l'opacitat dels sistemes d'intel·ligència artificial¹⁰²⁷. Mentre la màquina aprèn per ella mateixa -en el cas dels sistemes d'aprenentatge automàtic- i la intervenció humana està focalitzada en definir les tasques precises que ha de realitzar l'algoritme i les dades que s'usen, les persones no poden saber quin ha estat el procés de decisió, cada cop més basat en correlacions i deduccions que en causalitat. El raonament -artificial- de la màquina no és comparable al raonament -natural- de l'ésser humà, de manera que els resultats produïts no poden ser transparents ni explicables¹⁰²⁸. Per consegüent, s'estableix informació asimètrica entre aquelles persones que entren el sistema algorítmic i les persones sobre les quals es genera el resultat, ocasionant que els resultats no es puguin auditar, impedit possiblement la millor manera de ser conscient dels biaixos i les discriminacions¹⁰²⁹.

Per tant, el debat es desenvolupa en dos temps, en primer lloc si el dret a una explicació s'integra o no en l'RGPD¹⁰³⁰, i en segon lloc, si aquest dret és valuós quant a la informació que es rebrà tenint en compte que tot allò relatiu al tractament de dades mitjançant algorismes generalment és un procés invisible per la persona que proveeix les dades (ja que implica crear noves dades a través de la interrelació de les dades d'aquesta persona amb les dades d'altres).

¹⁰²⁵Conseil Constitutionnel, décision 2018-765DC, du 12 juin 2018, considerant 70. Disponible a: <https://www.conseil-constitutionnel.fr/decision/2018/2018765DC.htm> Darrera visita el 2 de setembre de 2023.

¹⁰²⁶GOODMAN, B., i FLAXMAN, S., "European Union Regulations on Algorithmic Decision-Making and a "Right to Explanation"", *AI Magazine*, vol. 38, núm. 3, 2017, pàg. 55.

¹⁰²⁷CASTETS-RENARD, C., "Accountability of Algorithms in the GDPR...", *op. cit.*, 2019, pàg. 101-108.

¹⁰²⁸*Idem, ibidem.*

¹⁰²⁹*Idem, ibidem.*

¹⁰³⁰ GERARDS, J. i ZUIDERVEEN, F., "Protected Grounds and the System...", *op. cit.*, 2022, pàg. 20. En aquest article s'exposa que les provisions de transparència específiques en referència a les decisions automatitzades rarament s'han aplicat a la pràctica.

Assumint que un dret a l'explicació s'integra en el marc de l'RGPD -de manera molt limitada donat que només existiria en el cas d'una decisió *únicament* automatitzada i dins el marc de la protecció de dades-, sembla inversemblant pretendre que tota persona que vegi les seves dades tractades dins d'un sistema massiu pugui obtenir una explicació adequada sobre la lògica del sistema arran de la qual pugui entendre que ha patit discriminació. En aquest sentit, un article crític amb el sistema algorítmic AMS, exposa que en el moment en què una persona que es veu avaluada per aquest sistema algorítmic busqui una explicació (per exemple, saber l'assignació específica d'un coeficient als seus atributs personals -el sistema algorítmic classifica les persones cercadores de feina en tres grups segons les seves oportunitats de trobar feina a curt, mig i llarg termini-), aquesta no podrà aclarir la multitud de factors que han contribuït a formar cada coeficient. No només això, sinó que és poc probable que el llenguatge tècnic i la documentació existent siguin suficients per explicar els detalls necessaris a qualsevol persona que no tingui formació prèvia en els camps de l'estadística, l'aprenentatge automàtic o les ciències de la computació¹⁰³¹. També, l'article 22 de l'RGPD sembla acotar un àmbit restringit del dret a l'explicabilitat, atès que no s'estén a decisions automatitzades dins el procés de decisió, inclús quan aquestes siguin especialment rellevants en el conjunt de factors que porten a la decisió final (per exemple, cribratge de currículums en un procés de selecció, finalment humanitzat). Per exemple, s'han presentat 1000 currículums, dels quals 15 han estat seleccionats per a l'examen humà, és a dir, que el gruix del biaix, on probablement radiqui la majoria de discriminacions per atributs protegits, haurà quedat inclòs en una fase íntegrament automatitzada, però que en haver-hi una persona al final del procés, no s'oferirà cap explicació¹⁰³².

Kaminski i Malgieri defensen que la transparència establerta a l'RGPD s'entén millor establint un sistema d'explicacions amb múltiples capes¹⁰³³. La transparència necessària per invocar drets no és la mateixa que es necessita per altres tipus de supervisió. La transparència individual està estructurada per ser comprensible i que es pugui exercir davant la justícia per part de persones individuals (art. 12 RGPD). Aquesta informació no es la mateixa que necessiten les persones expertes per valorar si el sistema està funcionant

¹⁰³¹ ALLHUTTER, D., CECH, F., FISCHER, F., GRILL, G. i MAGER, A., "Algorithmic Profiling of Job...", *op. cit.*, 2020, pàg. 14.

¹⁰³² RIVAS VALLEJO, P., "Sesgos de género en el uso...", *op. cit.*, 2022, pàg. 73.

¹⁰³³ KAMINSKI, M. E. i MALGIERI, G., "Algorithmic impact assessments...", *op. cit.*, 2021, pàg. 128 i 141. Es determinen les següents tipologies: explicacions individuals, explicacions grupals i explicacions sistèmiques, tant internes com externes.

correctament. Fer la transparència individual més útil per a les persones expertes en aquest àmbit la fa menys útil per a la resta d'individus. També s'afegeix el risc que si s'incrementa la rendició de comptes estructural afegint capes de supervisió de terceres parts, s'està distribuint informació individual a un major nombre de parts ¹⁰³⁴.

Rudin, molt crítica amb la idea de l'aprenentatge automàtic explicable, determina que els sistemes d'intel·ligència artificial s'anomenen caixes negres, ja que explicar una idea general de per què funcionen depèn més de la intuïció que no pas de la comprensió. Defensa que la creació de mètodes per explicar els models de caixes negres és una estratègia que continuarà perpetuant la mala praxi, fent-se palesa la necessitat de crear models que siguin inherentment explicables¹⁰³⁵. També mostren el seu escepticisme Edwards i Veale, en considerar que buscar un dret a l'explicació que es pugui exercir legalment pot ser en el millor dels casos una distracció i en el pitjor una manera de nodrir la fal·làcia de la transparència (en anglès *transparency fallacy*)¹⁰³⁶. Aquesta fal·làcia presumeix no només que els individus entenen els sistemes, sinó que també tenen maneres de discutir i debatre el seu significat, traslladant-los una càrrega enorme per buscar informació, interpretar-la i determinar el seu significat, és a dir, traslladant-los la responsabilitat d'entendre la decisió i la seva idoneïtat¹⁰³⁷.

Nogensmenys, per abordar la utilitat que pot tenir el dret a una explicació en el cas de les discriminacions cal fer dues distincions. En primer lloc, cal distingir que en el cas d'algoritmes no predictius (es limiten a automatitzar processos –p. e. revisió de requisits preestablerts en un conjunt de sol·licituds-) aquest dret pot ser útil en el sentit que pot aclarir certa informació de rellevància en la presa de decisió en operar en base a una norma, directriu o criteri preestablert; però en el cas dels algoritmes predictius, com *per se* hi ha un control molt més reduït de les raons per les quals el sistema actua com ho fa, una explicació sobre la seva lògica pot ser completament banal, ja que les causes de la

¹⁰³⁴ KAMINSKI, M.E., “Binary Governance...”, *op. cit.*, 2019, pàg. 1614.

¹⁰³⁵ RUDIN, C., “Stop explaining black box machine learning models for high stakes decisions and use interpretable models instead”, *Nature Machine Intelligence*, vol.1, núm. 5, 2019, pàg. 206-215. En aquest sentit, en l'àmbit de la salut, vegeu, per exemple: FUNDACIÓ TIC SALUT SOCIAL, “Guia sobre l'explicabilitat en la Intel·ligència Artificial”, 2022. Disponible a: <https://ticsalutsocial.cat/wp-content/uploads/2023/01/230226-Informe-Explicabilitat.pdf>. Darrera visita el 2 de setembre de 2023.

¹⁰³⁶ VEALE, M. i EDWARDS, L., “Slave to the Algorithm? Why...”, *op. cit.*, 2017, pàg. 81.

¹⁰³⁷ ANANNY, M., i CRAWFORD, K., “Seeing without Knowing: Limitations...”, *op. cit.*, 2018, pàg. 980.

discriminació es poden trobar en molts altres llocs a part dels criteris preestablerts pel funcionament del sistema (ens remetem al problema que acabem de relatar amb l'AMS).

En segon lloc, si entenguéssim que el dret a l'explicabilitat és útil tant en algorismes no predictius com predictius, interessa distingir entre la discriminació directa i la discriminació indirecta. En el cas de la discriminació directa (basada directament en un criteri discriminatori), el fet que la persona obtingui informació sobre la lògica emprada pel sistema -i la sàpiga interpretar-, pot suposar que pugui plantejar una evidència de la discriminació *prima facie* si el perfil elaborat s'ha basat explícitament en una categoria de dades sensible, podent-se produir llavors la inversió de la càrrega de la prova. En canvi, en la discriminació indirecta (basada en un criteri aparentment neutre però amb efectes discriminatoris) és molt més complicat que pugui ser útil, ja que malgrat individualment pugui donar-se certa informació, sempre serà necessària una altra perspectiva grupal, basada en proves estadístiques o protecció col·lectiva, ja que el perfil no es basarà explícitament en una categoria sensible de dades sinó en moltes dades interrelacionades, de manera que es necessitaran molts casos de persones discriminades en base al mateix criteri per establir discriminació *prima facie*.

Llavors, en el cas que es pugui demanar algun tipus d'explicació, ens preguntem quin seria el contingut òptim de la informació individual per donar a les persones que es veiessin subjectes a una decisió (únicament i no únicament) automatitzada per poder rebatre un resultat discriminatori.

3.2. Contingut de la informació individual necessària per iniciar un plet de discriminació

Com hem esmentat en el l'apartat 3.1 del capítol III, els actes de notificació de l'article 13 i 14 de l'RGPD s'apliquen abans del tractament de dades o de la presa de decisió i donen una descripció simple i genèrica de les activitats de tractament de dades dirigides a un públic general. L'article 15 proveeix d'informació similar a la que donen l'article 13 i 14, per consegüent, ofereixen un valor limitat per entendre i contestar les decisions. S'afegeixen els límits de l'article 15.4 i el considerant 63 de l'RGPD que protegeixen els secrets comercials o de propietat industrial -s'ha plantejat, així i tot, l'existència de

models de transparència que poden auditar i revelar informació rellevant per a la població de manera efectiva sense entrar en conflicte amb la propietat intel·lectual¹⁰³⁸. Tanmateix, creiem pertinent analitzar específicament l'article 13.2 (f), 14.2 (g) i 15.1 (h), on s'estableix que en el cas d'existir una decisió únicament automatitzada s'ha de proporcionar “*informació significativa sobre la lògica aplicada*”.

Considerem que el més òptim per poder fer front a la discriminació algorítmica seria una traducció de la lògica del sistema en conceptes que es puguin entendre i ser útils per a la persona que els rep. Per consegüent, hem trobat interessant la proposta, en essència, de “counterfactual explanations” (que traduirem com explicacions contrafàctiques) de Wachter, Mittelstadt i Russel¹⁰³⁹.

La idea que promouen és que en comptes de centrar-nos en explicar com les decisions s'han pres internament, es doni informació sobre quins fets externs poden ser diferents per poder arribar a l'objectiu desitjat¹⁰⁴⁰. Per exemple, en el cas de la denegació d'un préstec, una explicació contrafàctica donaria les raons per les quals s'ha donat una decisió particular (p. e. tenir un salari baix), proporcionaria les bases per contestar-la (p. e. la persona responsable del tractament ha emprat dades imprecises sobre el salari de la persona sol·licitant) i donaria recomanacions sobre com rebre els resultats volguts en el futur (p. e. si es tingués un salari de 4.000 euros més a l'any, el resultat podria haver estat positiu)¹⁰⁴¹.

Si extrapolem aquesta tipologia d'explicacions a casos en què l'algoritme hagi incorregut en discriminació, en la denegació d'un préstec, la raó particular seria (p. e. ets una dona), no hi hauria bases per contesta-la (ja que no hi ha dades imprecises sobre el fet de ser dona) i finalment, el resultat volgut en un futur perquè l'algoritme donés el resultat desitjat seria (p. e. no ser dona). Una decisió presa en base a aquests criteris seria susceptible de ser contestada. En definitiva, el que seria necessari saber, igual que en les decisions preses

¹⁰³⁸ DIAKOPOULOS, N., “Accountability in algorithmic decision making”, *Communications of the ACM*, vol. 59, núm. 2, 2016, pàg. 59.

¹⁰³⁹ WACHTER, S., MITTELSTADT, B., i RUSSEL, C., “Counterfactual explanations without...”, *op. cit.*, 2018, pàg. 841-887. Wachter, Mittelstadt i Russel plantegen que les explicacions contrafàctiques s'haurien de donar independentment que la decisió fos o no únicament automatitzada i dels seus efectes, abandonant les condicions i definicions estretes que imposa l'RGPD (tractament únicament automatitzat, tenir efectes legals o significativament similars i que no s'hagi introduït cap exempció (p. e. seguretat pública, l'execució de demandes civils o la seguretat nacional, entre d'altres) per part de l'estat membre (art. 23 RGPD)).

¹⁰⁴⁰ *Ídem*, pàg. 880.

¹⁰⁴¹ *Ídem*, pàg. 882.

per éssers humans, es concretaria en quins criteris han fonamentat la decisió, per poder, en cas de disconformitat, rebatre'ls.

En l'article de Wachter, Mittelstadt i Russel es presenten les explicacions contrafàctiques per tal que les persones puguin saber de quina manera poden canviar la decisió presa per l'algoritme i millorar el la seva situació personal. En canvi, a nosaltres, ens semblen d'interès atès que creiem que proporcionen la informació pertinent per saber en quins factors es basa l'algoritme per prendre la decisió, i com aquests factors poden revelar fàcilment un tracte discriminatori (potser en el cas de ser dona és molt evident, però podrien revelar, per exemple, que per obtenir una decisió favorable t'has de canviar de lloc de residència o de codi postal). També, perquè incorporen dades de funcionament de l'algoritme (quins factors ha valorat l'algoritme en la presa de decisió) en una explicació individual. Tot i que, com s'exposava en la crítica a l'AMS, s'ha de tenir en compte la probable impossibilitat de la persona per aclarir la multitud de factors que han contribuït a formar cada coeficient i que el llenguatge tècnic i la documentació existent siguin suficients per explicar els detalls necessaris a qualsevol persona que no tingui formació prèvia en els camps de l'estadística, l'aprenentatge automàtic o les ciències de la computació¹⁰⁴², creiem que s'ha de treballar en la direcció de donar resultats senzills i simples que puguin ser útils, i les explicacions contrafàctiques això ho poden proporcionar.

La proposta de Wachter, Mittelstadt i Russel sembla basar-se en el sistema creat als Estats Units d'Amèrica del Nord de l'Equal Credit Opportunity Act¹⁰⁴³ (d'ara en endavant, ECOA) i Fair Credit Reporting Act¹⁰⁴⁴ (d'ara en endavant, FCRA) -mitjançant la Regulation B¹⁰⁴⁵-, on, amb l'objectiu de prevenir la discriminació es dona l'oportunitat de saber on i com la sol·licitud de crèdit de la persona és deficient -per exemple, que ha residit poc temps en aquella àrea o que ha canviat recentment de feina¹⁰⁴⁶. En el cas de la notificació de la FCRA i l'EOCA, la persona només té accés al seu resultat específic i

¹⁰⁴² ALLHUTTER, D., CECH, F., FISCHER, F., GRILL, G. i MAGER, A., "Algorithmic Profiling of Job...", *op. cit.*, 2020, pàg. 14.

¹⁰⁴³ Equal Credit Opportunity Act, Pub. L. No. 93-495, 88 Stat. 1521 (1974) (codified as amended at 15 U.S.C. §§ 1691-1691f (2012)).

¹⁰⁴⁴ Fair Credit Reporting Act, Pub. L. No. 91-508, 84 Stat. 1127 (1970) (codified as amended at 15 U.S.C. §§ 1681-1681x (2012)).

¹⁰⁴⁵ Regulation B, 12 C.F.R. §§ 1002.1-16 (2018).

¹⁰⁴⁶ SELBST, A. D. i BAROCAS, S., "The Intuitive Appeal of Explainable Machines", *Fordham Law Review*, vol. 87, núm. 3, 2018, pàg. 1102.

per tant, no és possible observar un patró de discriminació indirecta, sinó únicament una explicació basada en el resultat.

Selbst i Barocas critiquen l'explicació basada en el resultat que, argumenten, serveix per responsabilitzar les persones que es veuen subjectes a una decisió sobre el resultat del sistema, en comptes de servir per demanar justificacions a aquelles que el dissenyen¹⁰⁴⁷. També la comparen amb l'explicació basada en la lògica del sistema de l'RGPD i consideren que aquesta última suposa dues millores en relació amb la perspectiva de l'explicació basada en els resultats. En primer lloc, una explicació basada en la lògica permet que qualsevol norma que sigui manifestament censurable esdevingui visible (contràriament a les explicacions en base als resultats que no aporten informació sobre les normes que determinen el sistema), i en segon lloc dona flexibilitat per, en alguns casos, revelar totes les normes darrere el procés de decisió, i en altres casos, només revelar-ne algunes, segons sigui necessari per entendre el funcionament del sistema¹⁰⁴⁸.

Si bé és cert que Selbst i Barocas tenen raó, en un camp tècnic que es presenta molt complex per qualsevol persona no especialitzada, considerem que la línia més efectiva a seguir actualment no és aquella que capti més matisos sinó aquella que porti uns mínims d'informació, és a dir, que porti senzillesa i utilitat¹⁰⁴⁹. Malgrat que les explicacions contrafàctiques no es promouen ni es troben emparades dins el marc de l'RGPD, el mecanisme que plantegen és rellevant en relació amb el contingut de la informació necessària per oposar-se a una decisió discriminatòria en el cas dels sistemes algorítmics, i és que es necessita tant informació sobre l'individu en qüestió (un resultat específic) com informació bàsica, entenedora i de funcionament general sobre l'algoritme (quins factors troba més rellevants l'algoritme per la presa de decisió). En les explicacions

¹⁰⁴⁷ *Ídem*, pàg. 1122. Expliquen els autors que aquest és un debat similar al de la notificació i el consentiment, on l'objectiu de la notificació és informar millor a les persones consumidores, assumint que millor informació portarà a resultats més preferibles. En general, exposen, això fracassa a protegir la privacitat de les persones, ja que construeix la privacitat com un assumpte de presa de decisió individual que una persona pot escollir a protegir, en comptes d'una cosa que es pot veure influenciada i afectada per altres amb més poder.

¹⁰⁴⁸ *Ídem*, pàg. 1108 i 1109.

¹⁰⁴⁹ En aquest sentit, vegeu: RIVAS VALLEJO, P., "Sesgos de género en el uso...", *op. cit.*, 2022, pàg. 65. L'autora es pregunta si realment cal saber com ha arribat a una conclusió el sistema de IA, o ens podem atendre en exclusiva al resultat, tenint en compte dues hipòtesis: la tesi de la conducta i la tesi del resultat. És a dir, què és més interessant per lluitar envers la discriminació, centrar-nos en com l'algoritme ha pres la decisió o en els efectes de l'algoritme? Tal com hem defensat al llarg d'aquest apartat, en aquest treball ens inclinem per determinar que serà més útil seguir la tesi del resultat, ja que és la tesi que és més assequible a persones que no tenen ni volen tenir cap formació respecte al funcionament dels sistemes algorítmics, i els dona, actualment, més possibilitats per desafiar-los.

contrafàctiques aquestes dues informacions hi són, i poden contrarestar-se, per tant, l'esforç que s'està fent és cap a generar un forma simple d'interactuar amb l'algoritme i detectar on és discriminatori.

Per consegüent, i en relació amb l'RGPD, en base a les explicacions contrafàctiques, els articles 13.2 (f), 14.2 (g) i 15.1 (h) poden ser útils quan es produeix discriminació algorítmica sota la condició que la “*informació significativa sobre la lògica aplicada*” a què es refereixen inclogui els factors que han determinat la decisió (codi postal, lloc de residència, sexe, salari, etc.). S'afegeix a aquesta condició que, en el cas de la discriminació indirecta, és necessari també un altre mecanisme (que abordarem seguidament) que aporti informació sobre la incidència de l'aplicació de l'algoritme en col·lectius oprimits (p. e. dones, persones amb discapacitat, persones més grans de 55 anys, etc.) que es pugui contrastar amb la informació individual obtinguda.

Cal subratllar que la presa de decisions mitjançant algoritmes ha de tenir les condicions adequades per entrar en l'àmbit d'aplicació de l'RGPD, ja que de no ser així, no es podrà comptar amb les seves proteccions. Així i tot, els resultats generats per dades d'entrenament esbiaixades violen el principi d'adequació de les dades (art. 5.1.d de l'RGPD) i la discriminació algorítmica no justificada també infringeix l'equitat en el tractament de dades (art. 5.1.a de l'RGPD)¹⁰⁵⁰, per tant, la discriminació algorítmica fàcilment pot entrar en l'àmbit d'aplicació de l'RGPD. Nogensmenys, també és cert que les disposicions de transparència de l'article 13.2 (f), 14.2 (g) i 15.1 (h) no s'apliquen si no es tracta de decisions *únicament* automatitzades. Això és un gran límit a la seva efectivitat i caldria estendre la seva aplicació a totes les decisions automatitzades.

Després d'examinar l'aportació que pot fer la possibilitat d'accés a informació individual que planteja l'RGPD per fer front a la discriminació algorítmica, volem examinar, tot seguit, un altre aspecte de la perspectiva de la transparència susceptible d'aportar protecció: la política de transparència que promou l'RGPD.

3.3. Política de transparència de l'RGPD

¹⁰⁵⁰ HACKER, P., “Teaching fairness to artificial intelligence...”, *op. cit.*, 2018, pàg. 1168.

La informació que l'RGPD permet proporcionar sempre és individual. Com ens referim a una protecció de dades personals, no hi ha cap article que plantegi revelar dades a alguna persona que no sigui en relació amb el seu cas propi. És evident que aquesta és la posició lògica dins una norma de protecció de dades personals. Tanmateix, en el cas de la discriminació algorítmica (sobretot indirecta) pot ser necessari saber de quina manera tracta el sistema algorítmic a col·lectius socialment oprimits¹⁰⁵¹ (és a dir, a diverses persones diferents que es troben agrupades per alguna raó). Com a mesures sistèmiques, l'RGPD planteja mesures de responsabilitat proactiva que integren elements de transparència, així que volem analitzar què pot oferir a la regulació de la discriminació algorítmica.

L'RGPD planteja la idea d'una política de la transparència que es podria dirigir a qui aplica els sistemes algorítmics per tal de treballar perquè siguin suficientment transparents com per determinar si estan essent discriminatoris o no. En conseqüència, considerem que té rellevància la política de transparència que promou l'RGPD (principi de transparència del tractament de l'article 5.1.a).

La protecció dels drets i llibertats de les persones físiques respecte al tractament de dades personals exigeix l'adopció de mesures tècniques i organitzatives apropiades amb la finalitat de garantir el compliment de l'RGPD¹⁰⁵². L'article 24 de l'RGPD, per exemple, llegit en base a l'article 1.2 de la norma, que estableix que l'RGPD protegeix drets i llibertats fonamentals de les persones entre ells el dret a la igualtat i a la no discriminació, genera el marc legal necessari que justifica la necessitat de les persones responsables de tractament d'introduir elements tècnics en els algoritmes per assegurar que no produeixen resultats discriminatoris¹⁰⁵³. Entre aquestes mesures -que ha d'adoptar la persona responsable del tractament-, es troba donar transparència a les funcions i al tractament de dades personals, permetent a les persones interessades supervisar el tractament de dades, així com a la persona responsable del tractament crear i millorar elements de seguretat¹⁰⁵⁴.

¹⁰⁵¹ En els assumptes Danfoss i Meister, el TJUE determina que el fet que la part demandada es negui a oferir la informació necessària per tal que les persones demandants puguin construir un cas pot justificar un canvi de la càrrega de la prova que forci a la part demandada a respondre, però considera necessaris indicis d'aquesta discriminació per poder operar el canvi del repartiment de la prova. Si fem una analogia amb un sistema algorítmic discriminatori, a no ser que la discriminació sigui directa, serà necessària informació sobre els efectes del sistema algorítmic en un col·lectiu oprimat determinat per poder aportar indicis de discriminació.

¹⁰⁵² Article 24, 28, 32 de l'RGPD, entre d'altres.

¹⁰⁵³ SORIANO ARNANZ, A., *Data Protection for the prevention...*, op. cit., 2021, pàg. 253.

¹⁰⁵⁴ Considerant 78 de l'RGPD.

S'instaura, amb aquest objectiu, per exemple, la figura de la persona delegada de protecció de dades com a eina de transparència¹⁰⁵⁵.

També, per exemple, aquesta política de transparència podria servir per identificar els algoritmes (és a dir, denominar i diferenciar cada sistema algorítmic de presa de decisió) i així, les persones que s'haguessin vist discriminades a través d'aquest, podent-lo identificar, tinguessin la possibilitat d'agrupar-se per desafiar les lògiques discriminatòries del sistema algorítmic. Com veurem en el següent capítol, la base de dades pública que instaura l'article 60 de la PRIA, pot dur a terme aquesta funció de transparència pública.

Nogensmenys, la transparència no sembla afectar a altres sistemes de responsabilitat proactiva, com són les auditories o les avaluacions d'impacte de protecció de dades¹⁰⁵⁶, a través de les quals es podria obtenir molta informació sobre el funcionament del tractament de dades, ja que l'RGPD no exigeix en cap moment que hi hagi una divulgació pública (o inclús a les persones interessades) dels seus resultats. En el mateix sentit, seria interessant anar més enllà, i plantejar les auditories i avaluacions d'impacte als sistemes algorítmics com un procés cíclic i públic que, entre altres coses, poguessin aportar estadístiques en relació amb els efectes dels algoritmes en grups protegits, forçant qui presumptivament ha discriminat a justificar que aquests efectes no tenen res a veure amb la discriminació¹⁰⁵⁷.

Com abordarem en el següent apartat relatiu a la responsabilitat proactiva, s'ha formulat la possibilitat que la divulgació al públic general dels resultats d'unes hipotètiques

¹⁰⁵⁵ L'article 37 de l'RGPD determina els tres casos en els quals és obligat designar un delegat de protecció de dades, a saber, si el tractament el duu a terme una autoritat pública o organisme públic, excepte els tribunals que actuïn en exercici de la funció judicial, si es tracta d'operacions de tractament que requereixin observació habitual i sistemàtica de persones interessades a gran escala o si es tracta d'un tractament a gran escala de categories especials de dades, o dades relatives a condemes o infraccions penals. A la doctrina espanyola, vegeu: SANTAMARÍA RAMOS, F. J., "La designación del delegado de protección de datos (Comentario al artículo 37 RGPD y al artículo 34 LOPDGDD)", a TRONCOSO REIGADA, A., (dir.) i VILLALBA CANO, L., (coord.), *Comentario al Reglamento General de Protección de Datos y a la Ley Orgánica de Protección de Datos Personales y Garantía de los Derechos Digitales*, Tomo I, Pamplona, Aranzadi, 2021, pàg. 2237-2265.

¹⁰⁵⁶ Les avaluacions d'impacte de protecció de dades no s'exigeixen en tots els tractaments de dades (art. 35.1 RGPD) ni s'exigeix que siguin públiques (art. 35.9 RGPD).

¹⁰⁵⁷ TJCE, sentència de 27 d'octubre de 1993, C-127/92, Dr. Pamela Mary Enderby contra Frenchay Health Authority and Secretary of State for Health, par. 19. S'estableix en aquest paràgraf que "*Cuando existen estadísticas significativas que ponen de manifiesto una diferencia considerable de retribución entre dos funciones del mismo valor; una de las cuales es desempeñada casi exclusivamente por mujeres y la otra fundamentalmente por hombres, el artículo 119 del Tratado exige que el empleador justifique dicha diferencia mediante factores objetivos y ajenos a cualquier discriminación por razón de sexo*".

avaluacions d'impacte algorítmiques (un pas més enllà de les avaluacions d'impacte de protecció de dades) poguessin proporcionar informació sobre els efectes dels sistemes algorítmics en col·lectius oprimits i servir com a documentació de contrast amb la informació individual.

Recapitulant, el dret a l'explicació pot ser útil si es produeix discriminació directa (atès que amb la informació individual obtinguda, la persona podria tenir suficients indicis com per presentar un cas *prima facie*), però si es produeix discriminació indirecta, malgrat que individualment pugui donar-se certa informació, sempre serà necessària una altra perspectiva grupal, basada en proves estadístiques o protecció col·lectiva. A més, la informació individual que es proporcioni hauria d'estar estructurada com a explicació contrafàctica, la tipologia més útil i senzilla per tal que una persona pugui detectar que s'ha vist discriminada. Quant als elements de transparència que integren les mesures de responsabilitat proactiva, podria ser útil, per exemple, si s'emprés per identificar sistemes algorítmics (és a dir, denominar i diferenciar cada sistema algorítmic de presa de decisió) i així, les persones que s'haguessin vist discriminades, poguessin identificar-los. Segons el nostre parer, la perspectiva a través de la qual l'RGPD tracta l'accés a la informació i la transparència i les eines que proporciona (ja siguin individuals, ja siguin mitjançant sistemes de transparència que imposa la responsabilitat proactiva) no són suficients en si mateixes per captar la discriminació algorítmica. Tot i així, assenten unes bases per poder crear un sistema que reguli i protegeixi les persones de les discriminacions algorítmiques que pugui emmirallar els punts clau de la transparència en l'RGPD, però que indefectiblement vagi més enllà. De la mateixa manera, en el cas d'altres eines de responsabilitat proactiva, com veurem en el següent apartat, cal anar més enllà per poder abordar la discriminació algorítmica.

4. Responsabilitat proactiva plantejada en l'RGPD

En el quart apartat del capítol tercer, hem analitzat quines eren les limitacions que trobàvem en els sistemes de gestió del risc plantejats en l'RGPD (avaluacions d'impacte de protecció de dades i auditories) per fer front a la discriminació algorítmica. En resum, hem conclòs que no estaven pensades per sistemes que evolucionessin amb el temps (com

els sistemes algorítmics de presa de decisió causants de discriminació) i que els paràmetres que els definien podien patir els mateixos biaixos que el tractament en si, amb el risc d'esdevenir processos merament formalistes i normatius. No obstant això, trobem en les eines de responsabilitat proactiva que planteja l'RGPD la voluntat de crear una política de compliment que generi en les pròpies organitzacions que empen sistemes algorítmics de presa de decisió consciència sobre l'impacte que poden tenir a la pràctica, i també, la necessitat de prendre mesures tant per superar estàndards de certificació com per crear sistemes més compromesos socialment.

Com ja hem esmentat en l'apartat anterior, la responsabilitat proactiva pot donar elements i obrir possibilitats per abordar la discriminació algorítmica. En primer lloc, pot clarificar en qui recau la responsabilitat d'una acció discriminatòria (4.1), en segon lloc, pot promocionar l'ús d'eines com els codis de conducta, les certificacions, els segells i les marques per incentivar la no-discriminació en els sistemes algorítmics (4.2), en tercer lloc, planteja la imposició de mesures des del disseny i per defecte (4.3), i en quart lloc, com a peça clau de les eines de responsabilitat proactiva que pot influir en la protecció envers la discriminació algorítmica, pot donar peu a la implementació d'avaluacions d'impacte algorítmiques, que s'han proposat com un pas més enllà de les avaluacions de protecció de dades (4.4).

4.1. Promoció d'una política del compliment i responsabilitat de l'acció discriminatòria

El principi de responsabilitat proactiva (concretat des de l'article 24 fins a l'article 43 de l'RGPD) estableix la necessitat d'incorporar una sèrie de garanties addicionals, més enllà d'un mínim, documentades i orientades a gestionar els riscos pels drets i llibertats dels individus, com per exemple l'obligació de mantenir un registre d'activitats de tractament (article 30 de l'RGPD)¹⁰⁵⁸.

La responsabilitat proactiva no és un principi nou. Es va recollir a les Directrius sobre la protecció de la privacitat i fluxos transfronterers de dades personals de l'Organització per

¹⁰⁵⁸ AEPD, *Adecuación al RGPD de tratamientos que incorporan Inteligencia Artificial. Una introducción*, Madrid, 2020, pàg. 20.

a la Cooperació i el Desenvolupament Econòmic (OCDE), efectives des del 23 de setembre de 1980¹⁰⁵⁹ i a la Directiva 95/46/CE. Els subjectes obligats (organitzacions que duen a terme tractament de dades personals) es posicionen com a mitjans necessaris per garantir el compliment de la normativa (han de buscar estratègies per lluitar envers la complexitat normativa i més en un món globalitzat) davant les autoritats de control i les persones interessades¹⁰⁶⁰. Com els sistemes d'intel·ligència artificial estan dissenyats per éssers humans, primerament s'ha d'identificar qui s'ha de considerar legalment responsable d'avaluar els riscos que poden ocasionar-se, monitoritzar el desenvolupament i complir amb el marc regulador vigent. En l'RGPD s'atorga responsabilitat a aquella persona que determina el sentit i la finalitat del tractament de les dades¹⁰⁶¹, concretament, la persona responsable de tractament.

Ha de ser capaç de demostrar mitjançant la responsabilitat proactiva que es respecten els principis que s'enumeren en l'article 5.1 de l'RGPD -tractament de manera lícita, lleial i transparent; amb finalitats determinades, explícites i legítimes sense un tractament ulterior per finalitats diferents; recollida de dades adequada, pertinent i limitada a la finalitat perseguida, ús de dades exactes i actualitzades (dret de la persona interessada a la rectificació i modificació); identificació de les persones interessades no més d'allò que sigui necessari; i garantia de la seguretat de les dades personals¹⁰⁶².

L'RGPD no especifica quins elements de responsabilitat proactiva cal implementar de forma obligatòria o la manera de fer-ho; simplement es limita a proporcionar un ampli grau de flexibilitat perquè tractaments de dades molt diferents s'adeqüin a la norma. Tot i així, estableix que les mesures s'han de seleccionar seguint un anàlisi basat en el risc pels drets i llibertats de les persones en relació amb el tractament de les seves dades personals i l'elaboració de perfils¹⁰⁶³.

¹⁰⁵⁹ OCDE, *Directrices sobre protección de la privacidad y flujos transfronterizos de datos personales*, París, 23 de septiembre de 1980.

¹⁰⁶⁰ DURÁN CARDO, B., "La responsabilidad del responsable del tratamiento y las obligaciones generales del responsable. El cumplimiento proactivo de la normativa de protección de datos: de la teoría a la práctica (Comentario al artículo 24 RGPD y al artículo 28 LOPDGDD)", a TRONCOSO REIGADA, A., (dir.) i VILLALBA CANO, L., (coord.), *Comentario al Reglamento General de Protección de Datos y a la Ley Orgánica de Protección de Datos Personales y Garantía de los Derechos Digitales*, Tomo I, Pamplona, Aranzadi, 2021, pàg. 1810.

¹⁰⁶¹ AMRAM, D., "The Role of the GDPR in Designing the European...", *op. cit.*, 2020, pàg. 8.

¹⁰⁶² Article 5.1 i 5.2 de l'RGPD.

¹⁰⁶³ AEPD, *Adecuación al RGPD de tratamientos que incorporan Inteligencia Artificial. Una introducción*, Madrid, 2020, pàg. 30.

Aquesta perspectiva basada en el risc (que inspira la proposta de regulació europea de la intel·ligència artificial, com veurem en el Capítol VI) té dues fases fonamentals: la primera és la identificació d'amenaques i l'avaluació del nivell de risc existent, i la segona és la gestió d'aquest risc mitjançant la implementació de mesures tècniques i organitzatives adequades i proporcionals per eliminar-lo o mitigar-lo, reduint l'impacte o la probabilitat que es materialitzin en les amenaces identificades, avaluant i controlant el risc residual, una vegada implementades les mesures seleccionades. S'ha d'identificar d'on procedeixen els riscos, si ho fan del tractament en si mateix (per exemple traslladar al sistema de decisions automatitzades perjudicis existents a la societat vinculats a categories especials de dades), o si es deriven del tractament (per exemple, del context social i dels efectes col·laterals relacionats amb l'objecte del tractament)¹⁰⁶⁴.

La persona responsable de dades té diverses obligacions que existeixen independentment de qualsevol acció del subjecte de dades. Aquest últim, segons estipula l'article 79 de l'RGPD, ha de tenir el dret a un recurs judicial efectiu quan consideri que els seus drets han estat violats com a resultat del tractament de les seves dades de manera no adequada a l'RGPD¹⁰⁶⁵. En canvi, els drets continguts en tercer capítol de l'RGPD només creen obligacions legals per a la persona responsable del tractament si el subjecte de dades activament escull exercir-los¹⁰⁶⁶.

Malgrat que la política de compliment la promogui la persona responsable de tractament, també és responsable si el tractament té efectes discriminatoris?

L'RGPD únicament fa referència a l'aspecte de la responsabilitat des del punt de vista de la protecció de dades, sense tenir en compte altres consideracions legals o ètiques que es puguin derivar de l'ús del sistema de decisió automatitzada¹⁰⁶⁷. Es considera que tota persona que hagi patit danys i perjudicis materials o immaterials com a conseqüència d'una infracció de l'RGPD tindrà dret a rebre per part de la persona responsable o delegada de tractament una indemnització per danys i perjudicis¹⁰⁶⁸. Es precisa que qualsevol responsable de tractament que hi hagi participat serà responsable en el cas que no s'hagin complert les obligacions establertes en l'RGPD dirigides específicament al

¹⁰⁶⁴ *Ídem*, pàg. 30.

¹⁰⁶⁵ Article 79 de l'RGPD.

¹⁰⁶⁶ WOLTERS, P.T.J., "The Control by and Rights of the Data Subject...", *op. cit.*, 2018, pàg. 7.

¹⁰⁶⁷ AEPD, *Adecuación al RGPD de tratamientos que incorporan Inteligencia Artificial. Una introducción*, Madrid, 2020, pàg. 19.

¹⁰⁶⁸ Article 82.1 de l'RGPD.

responsable, o hagi actuat al marge o en contra de les instruccions legals del responsable¹⁰⁶⁹. En el cas contrari, no serà responsable si demostra que no és responsable de cap manera del fet que ha causat els danys i perjudicis¹⁰⁷⁰. En el cas d'haver participat més d'un responsable, cada un serà considerat responsable de tots els danys i perjudicis amb la finalitat de garantir la indemnització efectiva de la persona interessada¹⁰⁷¹. Un/a responsable que hagi pagat la indemnització total podrà reclamar als/a les altres la part corresponent a la seva responsabilitat pels danys i perjudicis causats¹⁰⁷².

No obstant això, què passa quan existeixen biaixos implícits en l'algoritme que ocasionen la discriminació? És a dir, la persona responsable del tractament de dades ha dut a terme diligentment les seves tasques però l'algoritme que tracta les dades (per biaixos implícits, inconscients, històrics, etc.) genera resultats discriminatoris. Aquesta és una qüestió cabdal atès que s'ha de tenir en compte com s'empraran els sistemes algorítmics i els nous biaixos psicològics que se'n derivaran. Quan es tracta de sistemes sense supervisió, en el moment en què no s'adopta una posició crítica al respecte, el perill d'atribuir una autoritat o pes superior a una decisió presa per una eina automatitzada pot fer incrementar els riscos de la delegació de responsabilitat, ocasionant que, donada la situació en la qual s'hagi de reivindicar una discriminació, esdevingui complex identificar qui o què l'ha produïda¹⁰⁷³. No només pot esdevenir complex determinar qui o què l'ha produïda, sinó en quin moment s'ha produït, si causar discriminació ha estat l'objectiu del tractament o aquesta ha aparegut en algun moment del tractament¹⁰⁷⁴.

En referència als nous biaixos psicològics que es poden crear i el canvi en la significació que donem a conceptes determinats, DeCamp i Tilburt exposen que, encara que sigui amb bona intenció, parlar de confiança en relació a la intel·ligència artificial és un error de categorització, ja que s'assumeix que la intel·ligència artificial pertany a la categoria de coses a les quals se'ls hi pot depositar confiança. Defensen que confiar és posicionar alguna cosa de valor en responsabilitat d'un altre en circumstàncies de vulnerabilitat. Exemplifiquen que la relació ideal entre metge/essa i pacient reflecteix aquesta definició.

¹⁰⁶⁹ Article 82.2 de l'RGPD.

¹⁰⁷⁰ Article 82.3 de l'RGPD.

¹⁰⁷¹ Article 82.4 de l'RGPD.

¹⁰⁷² Article 82.5 de l'RGPD.

¹⁰⁷³ AEPD, *Adecuación al RGPD de tratamientos que incorporan Inteligencia Artificial. Una introducción*, Madrid, 2020, pàg. 49.

¹⁰⁷⁴ HUERTA ANGUIANO, J.A., ““Naturaleza Intrínseca”, “Contexto” o “Finalidad” en la...”, *op. cit.*, 2020, pàg. 10.

El/la pacient diposita la seva salut en la responsabilitat i discrecionalitat voluntària del/la metge/essa creient en la seva motivació ètica. La intel·ligència artificial, en canvi, no té voluntat, ni motius, ni caràcter¹⁰⁷⁵.

La responsabilitat és el principi segons el qual una persona o una organització legalment responsable d'un perjudici ha de donar una explicació o una compensació pel perjudici causat. La presa de decisions per part d'un algoritme ha creat diverses qüestions interessants en relació amb el caràcter intencional o no de la discriminació. Bertail, Bounie, Cléménçon i Waelbroeck afirmen que en el cas de la discriminació intencional, la responsabilitat, jurídicament, serà de la persona que programa o de l'organització que comercialitza l'algoritme. En canvi, en els casos en els quals els algoritmes puguin discriminar involuntàriament, la demostració de la prova és més complicada, i es pregunten si l'organització que comercialitza l'algoritme ha de ser responsable de la natura discriminatòria però no intencional del seu algoritme¹⁰⁷⁶.

Si no recau la responsabilitat d'un algoritme discriminatori però no intencional en l'organització que l'empra, no hi ha manera d'incentivar que les organitzacions facin recerca o inverteixin en suprimir els biaixos en els algoritmes i fer-los menys discriminatoris. Posem un exemple. Selbst apunta que malgrat qualsevol efecte desproporcionat en un grup social vulnerable, pot ser difícil guanyar un litigi contra una persona que contracta que decideix emprar un sistema completament automatitzat per predir la millor persona candidata. Considera que això passa per moltes raons, però principalment perquè la regulació antidiscriminació, de manera pràctica està basada en el defecte o l'error, de manera que l'impacte advers s'ha de poder atribuir a les accions de la persona que contracta¹⁰⁷⁷. No obstant això, i seguint la línia de raonament, si es prova que la persona o organització que contracta ha emprat un sistema algorítmic discriminatori (per tant, defectuós o que causa error), des de la nostra perspectiva - independentment de la coneixença que l'algoritme és discriminatori o no-, ha de ser responsable de l'acció discriminatòria.

¹⁰⁷⁵ DECAMP, M. i TILBURT, J. C., "Why we cannot trust artificial intelligence in medicine?" *The Lancet-Digital Health*, vol. 1, núm. 8, 2019, pàg. e-390.

¹⁰⁷⁶ BERTAIL, P., BOUNIE, D., CLÉMENÇON, S. i WAELBROECK, P., "Algorithms: biais, discrimination et équité", Télécom ParisTech, 2019, pàg. 18 i 19.

¹⁰⁷⁷ SELBST, A. D., "An institutional view of Algorithmic Impact Assessments", *op. cit.*, 2021, pàg. 129 i 130.

Cal esmentar que hi ha propostes d'interès que fonamenten la responsabilitat d'aquella persona que pren la decisió en un cas de discriminació algorítmica per la seva conducta imprudent. S'incorreria en imprudència o negligència, en cas d'adoptar decisions fundades en sistemes algorítmics amb resultats esbiaixats, donada la possibilitat de prevenir i rectificar biaixos i estereotips. Ja que l'evolució tecnològica permet saber que existeix el risc de biaixos implícits, poden utilitzar algoritmes predictius per reduir-ne el risc, per consegüent, davant una manca de prevenció han de ser responsables de discriminació –intencional-, derivada de la seva negligència o imprudència¹⁰⁷⁸. En aquest sentit, interessa com la Llei 15/2022 estableix que la persona física o jurídica que causi discriminació repararà el dany proporcionant una indemnització i restituint a la víctima en la situació anterior a l'incident discriminatori, quan sigui possible, i seran igualment responsables del dany causat les persones que contracten o prestadores de serveis quan la discriminació es produeixi en el seu àmbit d'organització i direcció i no hagin aplicat mètodes i instruments per la detecció, aplicació de mesures preventives i l'articulació de mesures adequades pel cessament de situacions discriminatòries¹⁰⁷⁹. També resulta interessant l'article 44 de la Llei 19/2020, que es titula “Responsabilitat”, i estableix que són responsables de les infraccions tipificades en la llei les persones físiques o jurídiques, públiques o privades, que per acció o omissió, incorren en els supòsits d'infracció, i si una infracció és imputable a diverses persones i no és possible determinar el grau de participació de cadascuna en la comissió de la infracció, n'han de respondre solidàriament¹⁰⁸⁰. Aquesta última apreciació és interessant en el cas que no es pogués determinar si la responsabilitat és del proveïdor o de l'empresa o organització que utilitza el sistema algorítmic, ja que al no poder-se determinar el grau de participació de cadascuna, haurien de respondre solidàriament.

Resumidament, la persona responsable de tractament promou una política de compliment de l'RGPD. La responsabilitat, independentment de qui hagi creat el sistema algorítmic, llevat en el cas que la persona responsable del tractament hagi causat per falta de diligència el resultat discriminatori, sembla recaure en l'organització que el posa en pràctica, de la mateixa manera que és la persona responsable de tractament qui promou un política de compliment de l'RGPD, però qui finalment se'n beneficia no és la persona

¹⁰⁷⁸SÁEZ LARA, C., “El algoritmo como protagonista de la relación laboral...”, *op. cit.*, 2020, pàg. 51; PÁEZ, A., “Negligent Algorithmic Discrimination”, *op. cit.*, 2021, pàg. 20.

¹⁰⁷⁹ Article 27 en relació amb l'article 25.1 de la Llei 15/2022.

¹⁰⁸⁰ Article 44 Llei 19/2020.

en si, sinó l'organització per a la qual treballa. Sembla que el més raonable és que l'organització que empra l'algoritme ha d'assegurar-se que aquest no és discriminatori, i si es prova que ho és, ha d'assumir-ne la responsabilitat.

Tot seguit abordem en què pot beneficiar l'existència de codis de conducta, certificacions, segells i marques per regular i disminuir la discriminació causada per sistemes algorítmics.

4.2. Sistemes d'autoregulació: Codis de conducta, certificacions, segells i marques

Per instaurar aquesta política de compliment, a part de la figura de la persona responsable del tractament o la persona delegada de protecció de dades, també es proposen mecanismes com els codis de conducta, les certificacions, els segells i les marques. Creiem rellevant analitzar quin impacte poden tenir aquests mecanismes en la protecció envers la discriminació.

L'objectiu és facilitar i crear una cultura efectiva del compliment de l'RGPD a partir de l'autoregulació¹⁰⁸¹ duta a terme per les entitats, institucions i organismes competents, partint de la base que la norma vigent en matèria de protecció de dades, els atorga una naturalesa especial dins el món jurídic. L'RGPD planteja la diligència de les persones obligades com una peça clau del sistema de compliment¹⁰⁸². En el context d'aquesta cultura del compliment ha d'introduir-se el compromís de treballar per reduir al màxim possible la discriminació dels sistemes algorítmics emprats.

¹⁰⁸¹ COMMISSION OF THE EUROPEAN COMMUNITIES, *First Report on the Implementation of the Data Protection Directive (95/46/EC)*, COM (2003). 265 final, Brussels, 15 may 2003, pàg. 26. Ja l'any 2003 la Comissió de les Comunitats Europees va exposar que estava decebuda que poques organitzacions haguessin plantejat codis de conducta sectorials, encoratjant als sectors i grups d'interès a prendre un rol molt més proactiu, ja que creia que l'autoregulació, i particularment els codis de conducta, haurien de jugar un rol important en el desenvolupament de la protecció de dades tant a la UE com a l'exterior, sobretot per evitar legislació excessivament detallada.

¹⁰⁸² PRIETO HERGUETA, J., "Los códigos de conducta (Comentario a los artículos 40 y 41 RGPD y a los artículos 38 y Disposición transitoria segunda LOPDGDD)", a TRONCOSO REIGADA, A., (dir.) i VILLALBA CANO, L., (coord.), *Comentario al Reglamento General de Protección de Datos y a la Ley Orgánica de Protección de Datos Personales y Garantía de los Derechos Digitales*, Tomo I, Pamplona, Aranzadi, 2021, pàg. 2363.

L'RGPD promou la creació de codis de conducta (es proposen incentius per crear-ne en els articles 28.5, 32.3, 35.8, 46.2 (e) i 83.1 (j)). Un codi de conducta és un document escrit i de compliment obligatori per al sector que el subscriu i que recull bones pràctiques per desenvolupar i complir en el tractament de dades personals. Dos elements es poden destacar en relació amb la seva naturalesa: la capacitat autoreguladora de les parts i el caràcter vinculant del codi. No es tracta tant d'atribuir una denominació jurídica al codi, que no és una llei, ni una font del dret¹⁰⁸³, sinó de posicionar-lo en un esquema de normes aplicables, que s'ha de realitzar analitzant les seves característiques essencials, que l'aproximen al món d'allò exigible i per tant de la normativitat¹⁰⁸⁴. Així i tot, ni l'RGPD ni la LOPDGDD imposen cap obligació d'elaboració¹⁰⁸⁵.

Els codis de conducta (com també els sistemes de certificació, segells i marques) es promouen per part del Consell Europeu, la Comissió Europea i les Autoritats de Protecció de Dades i requereixen d'organismes de control i supervisió. Han de ser aprovats per les autoritats de control, el comitè o la comissió, i la supervisió del seu compliment podrà ser realitzada per un organisme independent que hagi estat acreditat per a aquesta finalitat per l'autoritat de control competent. La seva efectivitat depèn, en gran mesura, de què organismes supervisors (art. 41.1 de l'RGPD) tinguin el nivell adequat de perícia en relació amb l'objecte del codi i que hagin estat acreditats per a la finalitat concreta per l'autoritat de control competent. És necessari que els codis de conducta incloguin un regim de sancions, sinó l'òrgan supervisor no tindrà eines adequades per obligar al compliment.

D'una banda, segons certa part de la doctrina, els codis de conducta tal com es recullen a l'RGPD, s'obren pas -gairebé es pot dir que es consoliden-, com una eina complementària que apuntala les normes i ajuda a garantir el seu compliment¹⁰⁸⁶. En canvi, d'altra banda,

¹⁰⁸³ OBSERVATORI D'ÈTICA EN INTEL·LIGÈNCIA ARTIFICIAL DE CATALUNYA (OEIAC), *Intel·ligència artificial, ètica i societat...*, op. cit., 2021, pàg. 8. Fent una analogia amb els codis d'ètica, s'estableix en l'informe referenciat que l'eina que ha proliferat més en el camp de l'ètica de la IA actualment és el codi d'ètica que busca guiar el disseny i desplegament de sistemes d'IA, però està limitat en aspectes clau, ja que manca d'un marc universalment acordat, i no és vinculant com una legislació específica.

¹⁰⁸⁴ SERRANO PÉREZ, M. M., "Los códigos de conducta (Comentario artículo 40 RGPD y al artículo 38 y Disposición transitoria segunda LOPDGDD)", a TRONCOSO REIGADA, A., (dir.) i VILLALBA CANO, L., (coord.), *Comentario al Reglamento General de Protección de Datos y a la Ley Orgánica de Protección de Datos Personales y Garantía de los Derechos Digitales*, Tomo I, Pamplona, Aranzadi, 2021, pàg. 2385 i 2386.

¹⁰⁸⁵ PRIETO HERGUETA, J., "Los códigos de conducta...", op. cit., 2021, pàg. 2364.

¹⁰⁸⁶ CAZURRO BARAHONA, V., "La supervisión de los códigos de conducta (Comentario al artículo 41 RGPD)", a TRONCOSO REIGADA, A., (dir.) i VILLALBA CANO, L., (coord.), *Comentario al*

s'apunta que sobre la base normativa anterior a l'RGPD (Directiva 95/46/CE), com es pot apreciar per l'escàs nombre de codis de conducta inscrits en el Registre de l'Agència Espanyola de Protecció de Dades, no han tingut una gran acceptació, malgrat que s'impulsés el seu desenvolupament i es destaqués la importància de l'autoregulació per augmentar el grau de coneixement i implantació del dret fonamental a la protecció de dades¹⁰⁸⁷.

L'RGPD també promou sistemes de certificació (art. 42). A través dels mecanismes de certificació, segells i marques, les persones responsables del tractament poden demostrar el compliment de les seves respectives obligacions i es permet a les persones interessades avaluar el nivell de protecció de dades de les organitzacions. Els mecanismes de certificació de protecció de dades suposen la validació d'un sistema de gestió, però l'obtenció d'una certificació¹⁰⁸⁸ no eximeix de la responsabilitat derivada de l'incompliment de les exigències normatives¹⁰⁸⁹, i l'RGPD planteja sancions per les infraccions de certificació. Amb independència de l'èxit del procés certificador, l'organització no podrà negligir el compliment dels procediments que van donar lloc a l'obtenció de la certificació, segell o marca, doncs, la pròpia entitat certificadora i/o autoritat de control, haurà de dur a terme una supervisió que podrà donar lloc a la retirada i a la imposició de sancions per incompliment¹⁰⁹⁰. Per tant, no es confiarà únicament en les manifestacions de fabricants i distribuïdors sinó que s'hauran d'establir mecanismes basats en tercers confiàbles per determinar els nivells necessaris de qualitat¹⁰⁹¹. L'article 42.7 de l'RGPD estableix que una certificació únicament és vàlida per un període màxim de tres anys abans de ser renovada.

Reglamento General de Protección de Datos y a la Ley Orgánica de Protección de Datos Personales y Garantía de los Derechos Digitales, Tomo I, Pamplona, Aranzadi, 2021, pàg. 2419.

¹⁰⁸⁷ PRIETO HERGUETA, J., "Los códigos de conducta...", *op. cit.*, 2021, pàg. 2376.

¹⁰⁸⁸ L'organisme de certificació a espanya és l'Entidad Nacional de Acreditación (ENAC). En relació als organismes de certificació, a la doctrina espanyola, vegeu: SEMPERE SAMANIEGO, J., "Organismo de certificación (Comentario al artículo 43 RGPD y al artículo 39 LOPDGD)", a TRONCOSO REIGADA, A., (dir.) i VILLALBA CANO, L., (coord.), *Comentario al Reglamento General de Protección de Datos y a la Ley Orgánica de Protección de Datos Personales y Garantía de los Derechos Digitales*, Tomo I, Pamplona, Aranzadi, 2021, pàg. 2443-2456.

¹⁰⁸⁹ HELGUERO SÁINZ, J. i PARDO DE VERA, M. M., "Mecanismos de certificación en la nueva normativa de protección de datos de carácter personal (Comentario al artículo 42 RGPD y al artículo 39 LOPDGD)", a TRONCOSO REIGADA, A., (dir.) i VILLALBA CANO, L., (coord.), *Comentario al Reglamento General de Protección de Datos y a la Ley Orgánica de Protección de Datos Personales y Garantía de los Derechos Digitales*, Tomo I, Pamplona, Aranzadi, 2021, pàg. 2424.

¹⁰⁹⁰ *Ídem*, pàg. 2441.

¹⁰⁹¹ AEPD, *Adecuación al RGPD de tratamientos que incorporan Inteligencia Artificial. Una introducción*, Madrid, 2020, pàg. 34.

D'una banda, segons les 23 entrevistes a persones expertes i/o interessades en el desenvolupament i impacte de la intel·ligència artificial procedents de diferents àmbits, des de l'acadèmic a l'industrial, passant per l'administració pública i la ciutadania, fetes per part de l'Observatori d'ètica en intel·ligència artificial de Catalunya entre els mesos de febrer i abril de 2021, s'exposa que *“sobre el tema de si és possible que els algorismes utilitzats en els sistemes d'IA siguin justos, sobretot quan són propietat privada de corporacions i no són accessibles per al control públic, les posicions són majoritàriament optimistes i denoten la importància de desenvolupar sistemes d'IA en els quals pugui integrar-se un principi de justícia social. Per a això, s'apunta que el més factible és utilitzar una certificació que permeti que les dades i el disseny de l'algorisme no estan esbiaixats i que existeix la possibilitat d'una rendició de comptes o auditoria del sistema d'IA per garantir-ho”*¹⁰⁹². En aquest sentit, interessa l'article 23 de la Llei 15/2022, relatiu a la intel·ligència artificial i mecanismes de presa de decisió automatitzats, que en el seu apartat quart afirma que es promourà un segell de qualitat dels algorismes¹⁰⁹³.

Tanmateix, d'altra banda, s'alerta per part d' Edwards i Veale que no hi ha proves que les persones vegin els segells i les certificacions com a indicadors de confiança, fet que pot indicar que les organitzacions no estan disposades a pagar i fer un esforç per fer servir aquests mecanismes, a no ser que en fer-ho puguin evitar una regulació més exhaustiva¹⁰⁹⁴.

El més probable en el cas de la discriminació serà que sigui un element a introduir en els codis de conducta i un factor que es tingui en compte pels organismes que donen certificacions, segells i marques. Si es té en compte en aquest sentit, s'haurà de verificar periòdicament i cíclicament el compliment del sistema algorítmic dels estàndards de certificació (quan es vol implantar, avaluant-se el sistema després d'haver-se implantat i arran de cada avaluació millorant-se si escau, i així cíclicament), atès que hi ha risc que es produeixi una evolució amb el temps i que arran d'aquesta, el sistema ja no compleixi amb els estàndards.

¹⁰⁹² OBSERVATORI D'ÈTICA EN INTEL·LIGÈNCIA ARTIFICIAL DE CATALUNYA (OEIAC), *Intel·ligència artificial, ètica i societat...*, op. cit., 2021, pàg. 132.

¹⁰⁹³ Article 23.4 de la Llei 15/2022.

¹⁰⁹⁴ EDWARDS, L. i VEALE, M., “Enslaving the Algorithm: From a “Right to an Explanation” to a “Right to Better Decisions”?”, *IEEE Security and Privacy*, vol. 16, núm. 3, 2018, pàg. 52.

A continuació, estudiem l'obligació de protecció de dades des del disseny i per defecte i la seva incidència en la mitigació dels efectes discriminatoris dels sistemes algorítmics.

4.3. Igualtat de tractament des del disseny i per defecte

L'article 25 de l'RGPD fa referència a la protecció de dades des del disseny i per defecte, un concepte derivat de la privacitat des del disseny, que sempre ha format part de la protecció de dades, i és un element clau de la perspectiva de gestió de riscos de l'RGPD. En el primer apartat s'exigeix a les persones responsables del tractament, que adoptin tant en el moment de determinar els mitjans de tractament com en el moment del tractament, fins que aquest ja no és operatiu (és a dir, en tot el cicle de vida del sistema), mesures tècniques i organitzacionals apropiades per aplicar de forma efectiva els principis de protecció de dades i integrar les garanties necessàries en el tractament per complir amb l'RGPD i protegir els drets de les persones interessades¹⁰⁹⁵ (adreçant els possibles riscos del tractament¹⁰⁹⁶).

En el segon apartat de l'article, s'especifica que la persona responsable de tractament haurà d'implementar mesures tècniques i organitzatives adequades per garantir que per defecte només siguin objecte de tractament les dades personals que siguin necessàries per a les finalitats específiques del tractament¹⁰⁹⁷. Aquest apartat està estretament vinculat amb els principis de limitació de la finalitat i minimització de dades, i suposa que s'han d'especificar les dades personals necessàries abans d'iniciar el tractament i informar de manera apropiada a les persones que només es tractaran les dades necessàries per la finalitat específica.

Cal apuntar que el principi de minimització de dades (només tractar aquelles dades rellevants per la finalitat del tractament) està en tensió amb les tecnologies de l'anàlisi massiva de dades (per exemple, l'aprenentatge automàtic), atès que generen millors resultats quan tracten més quantitat de dades. Considerant que les persones responsables

¹⁰⁹⁵ Article 25.1 de l'RGPD.

¹⁰⁹⁶El considerant 75 de l'RGPD, com hem apuntat, enumera els riscos del tractament i un d'ells és la discriminació.

¹⁰⁹⁷Article 25.2 de l'RGPD.

del tractament tendeixen a formular les finalitats del tractament de manera genèrica i que és molt difícil controlar totes les dades que s'introdueixen, no sembla un principi que suposi moltes limitacions a les persones responsables del tractament¹⁰⁹⁸.

Finalment, en el tercer apartat de l'article, es precisa que es podrà utilitzar un mecanisme de certificació aprovat sota els criteris que determina l'article 42 de l'RGPD per acreditar el compliment de les obligacions de protecció de dades des del disseny i per defecte¹⁰⁹⁹. Especifica el Comitè Europeu de Protecció de Dades en les "Directrius 4/2019 sobre l'article 25 de l'RGPD" que, tot i que s'obtingui una certificació d'acord amb l'article 42 de l'RGPD, la persona responsable del tractament té l'obligació de monitoritzar contínuament i seguir millorant la rendició de comptes¹¹⁰⁰.

El considerant 78 de l'RGPD precisa que l'adopció de polítiques internes i l'aplicació de mesures que compleixin els principis de protecció de dades des del disseny i per defecte s'han de dur a terme per poder demostrar la conformitat del tractament amb l'RGPD, podent consistir les mesures en qüestió, entre d'altres, en reduir al màxim el tractament de dades personals, pseudonimitzar-les el més aviat possible i donar transparència a les funcions i al seu tractament, possibilitant a les persones interessades la supervisió del tractament de dades i a la persona responsable de tractament la creació i millora d'elements de seguretat¹¹⁰¹.

El funcionament de la protecció de dades des del disseny i per defecte implica que la persona responsable del tractament consideri les activitats de tractament, els riscos que poden causar a les persones (p. e. discriminació), i les mesures disponibles per assegurar el compliment dels principis de protecció de dades i la protecció de drets individuals, possibilitant l'obtenció d'una certificació que ajudi a demostrar que s'està complint i implementant la protecció de dades des del disseny i per defecte.

Els elements clau de l'equitat en la protecció de dades des del disseny i per defecte han d'incloure la no discriminació¹¹⁰². Per tant, és de l'article 25 de l'RGPD que deriva el

¹⁰⁹⁸ SORIANO ARNANZ, A., *Data Protection for the prevention...*, op. cit., 2021, pàg. 122.

¹⁰⁹⁹ Article 25.3 de l'RGPD.

¹¹⁰⁰ EUROPEAN DATA PROTECTION BOARD, *Guidelines 4/2019 on Article 25 Data Protection by Design and by Default*, Brussels, 20 October 2020, pàg. 29.

¹¹⁰¹ Considerant 78 de l'RGPD.

¹¹⁰² EUROPEAN DATA PROTECTION BOARD, *Guidelines 4/2019 on Article 25 Data Protection...*, op. cit., pàg. 18.

concepte “igualtat des del disseny” que part de la doctrina considera necessari que s’apliqui al desenvolupament dels algoritmes com a estàndard¹¹⁰³ i que suposa estructurar els algoritmes de manera que siguin sensibles a les formes predominants de discriminació, tenint en consideració l’existència d’estructures socials de discriminació, intentant detectar discriminació, corregir-la i si escau incorporar nocions de discriminació sistèmica per corregir desequilibris socials entre membres de grups avantatjats i desavantatjats¹¹⁰⁴.

En aquest sentit, Diver critica que l’anàlisi legal sobre els efectes normatius del codi font dels sistemes algorítmics s’hagi centrat en els factors humans, en la decisió d’usar el codi font per implementar una norma, i no en les característiques materials del disseny del codi font. Defensa que davant la qüestió sobre com assegurar que el codi font respecti la normativa vigent, donada la natura privatitzada dels ambients en els quals el codi font es produeix, s’hauria d’apostar per imposar els límits constitucionals sobre quin tipus de codi hauria de produir-se, per una normativa implementada en el moment de producció del codi font que defineixi els seus possibles usos o aclareixi com s’hauria d’usar¹¹⁰⁵.

Sembla un bon mecanisme per incentivar les organitzacions a treballar per disminuir les possibilitats de resultats discriminatoris en els sistemes algorítmics que emprin. Cotino Hueso planteja que no seria estrany considerar la il·legalitat d’un sistema d’IA per la inacció o l’incompliment dels deures preventius de les persones responsables del seu ús en el disseny, control o avaluació del sistema, ja que l’article 4.1 en relació amb l’article

¹¹⁰³ Vegeu, per exemple: YANISKY-RAVID, S. i HALLISEY, S., “Equality and privacy by design: a new model of artificial intelligence data transparency via auditing, certification and safe harbor regimes”, *Fordham Urban Law Journal*, vol. 46, núm. 2, 2019, pàg. 428-486; HACKER, P., “Teaching fairness to artificial intelligence...”, *op. cit.*, 2018, pàg. 1146; XENIDIS, R. i SENDEN, L., “EU Non-discrimination law...”, *op. cit.*, 2020, pàg. 179.

¹¹⁰⁴ SORIANO ARNANZ, A., *Data Protection for the prevention...*, *op. cit.*, 2021, pàg. 199-202.

¹¹⁰⁵ DIVER, L., “Digisprudence: the design of legitimate code”, *Law, Innovation and Technology*, vol. 13, núm. 2, 2021, pàg. 352-353.

25.1 de la Llei 15/2022¹¹⁰⁶ ho disposa d'aquesta manera¹¹⁰⁷.

Nogensmenys, cal referenciar certes limitacions amb les quals aquest mecanisme de protecció pot trobar-se. Les decisions automatitzades, com ja hem vist, presenten riscos específics, per exemple, aquells derivats de l'execució del sistema (p. e. errors i biaixos) o aquells derivats de la manca d'informació i transparència sobre el procés decisió. Aquestes particularitats, suposen que la protecció des del disseny i per defecte sigui més complicada d'implementar. Atesa l'opacitat de certs sistemes algorítmics, és possible que els responsables del tractament no tinguin la informació necessària per posar en pràctica elements per protegir la discriminació, esdevenint una norma buida¹¹⁰⁸, especialment si aquelles persones que han d'aplicar la norma, no tenen els recursos per determinar clarament quines proteccions són necessàries en cada cas.

També, cal apuntar que sovint existeixen conflictes de valors que dificulten la identificació i implantació dels mitjans necessaris per a una protecció efectiva. Per exemple, la protecció de dades contempla un gran ventall de drets per exercir que poden col·lidir entre ells¹¹⁰⁹. La protecció des del disseny es pot veure, endemés, restringida per limitacions del disseny com a pràctica. L'ús del terme “disseny”, sovint submergeix una multitud heterogènia de visions sota un conjunt de pràctiques relativament estret, on les forces dominants sovint minimitzen i esborren el treball de les minories¹¹¹⁰. Sota aquestes circumstàncies, un instrument de protecció pot acabar reforçant les estructures de poder existents i creant o exposant diverses formes de vulnerabilitat. També s'ha criticat que l'article 25 de l'RGPD apel·la a eines que s'han revelat com absolutament insuficients en el camp de les inferències per aprenentatge automàtic (la pseudonimització i altres

¹¹⁰⁶ Article 4.1 de la Llei 15/2022 disposa que “*queda prohibida toda disposición, conducta, acto, criterio o práctica que atente contra el derecho a la igualdad. Se consideran vulneraciones de este derecho la discriminación, directa o indirecta, por asociación y por error; la discriminación múltiple o interseccional, la denegación de ajustes razonables, el acoso, la inducción, orden o instrucción de discriminar o de cometer una acción de intolerancia, las represalias o el incumplimiento de las medidas de acción positiva derivadas de obligaciones normativas o convencionales, la inacción, dejación de funciones, o incumplimiento de deberes*”. L'article 25.1 de la Llei 15/2022 disposa que “*La protección frente a la discriminación obliga a la aplicación de métodos o instrumentos suficientes para su detección, la adopción de medidas preventivas, y la articulación de medidas adecuadas para el cese de las situaciones discriminatorias*”.

¹¹⁰⁷ COTINO HUESO, L., “Los usos de la inteligencia artificial en el sector público, su variable impacto y categorización jurídica”, 2023, *Revista Canaria de Administración Pública*, núm. 1, 2023, pàg. 230.

¹¹⁰⁸ TAMO-LARRIEUX, A., *Designing for Privacy and Its Legal Framework: Data Protection by Design and Default for the Internet of Things*, Cham, Springer, 2018, pàg. 209.

¹¹⁰⁹ VEALE, M., BINNS, R., i AUSLOOS, J. “When Data Protection by Design and Data Subject Rights Clash”, *International Data Privacy Law*, vol. 8, núm. 2, 2018, pàg. 121.

¹¹¹⁰ BENJAMIN, R., *Race After Technology: Abolitionist Tools for the New Jim Code*, Medford, Polity, 2019, pàg. 176.

tècniques a les que fa referència el precepte per anonimitzar dades o de despullar-les dels trets personals), atès que qualsevol tipus d'atribut és susceptible de rastreig o inferència¹¹¹¹. Si bé és cert que existeixen aquestes limitacions, sembla una protecció que pot tenir recorregut per incentivar que les organitzacions dirigeixin esforços a treballar en reduir la possibilitat de discriminació en els seus sistemes algorítmics.

La protecció de dades des del disseny i per defecte és un concepte que integra moltes mesures per abordar riscos del tractament de dades i, entre elles, es troben les avaluacions d'impacte de protecció de dades¹¹¹². Tot seguit presentem les avaluacions d'impacte algorítmiques, un pas més enllà de les avaluacions d'impacte de protecció de dades, estudiant com poden participar en fer els sistemes algorítmics menys discriminatoris.

4.4. De les avaluacions d'impacte de protecció de dades a les avaluacions d'impacte algorítmiques

En aquest apartat analitzem les avaluacions d'impacte algorítmiques, que s'han defensat per part de la doctrina com a una millor forma de mitigar els riscos envers els drets fonamentals dels sistemes algorítmics que les avaluacions d'impacte de protecció de dades, que veuen la seva efectivitat limitada (com hem estudiat en l'apartat 4.1 del capítol III), ja que només s'apliquen a sistemes de risc alt, no s'exigeix que siguin públiques, analitzen l'impacte del sistema de manera inicial sense comptar amb la possible evolució del sistema i poden patir, en molts casos, els mateixos biaixos implícits que pateix el sistema algorítmic en si. Ens preguntem, llavors, quina incidència poden tenir les proposades avaluacions d'impacte algorítmiques en la mitigació de la discriminació (4.4.1) i els límits d'aquesta proposta (4.4.2), analitzant finalment la possibilitat que aquestes puguin crear documentació sobre la incidència del sistema algorítmic en col·lectius de població determinats que generin ulteriorment prova per un cas de discriminació *prima facie* (4.4.3).

¹¹¹¹ RIVAS VALLEJO, P., “Sesgos de género en el uso...”, *op. cit.*, 2022, pàg. 71.

¹¹¹² MICHELAKAKI, C. i BARROS VALE, S., *Unlocking Data Protection By Design & By Default: Lessons from the Enforcement of Article 25 GDPR*, Future of Privacy Forum, Washington, DC, 19 May 2023, pàg. 61.

4.4.1. Propostes d'avaluacions d'impacte algorítmiques

Es planteja una avaluació d'impacte algorítmica com un procés en el qual una persona desenvolupadora d'un sistema algorítmic vol anticipar, avaluar i investigar els danys potencials del sistema abans de la seva implementació, documentar les troballes que faci i inclús publicar-les o informar a una persona o entitat reguladora¹¹¹³. Hauria de ser una acció continuada: un procés que produeix resultats i informes però que a la vegada inclou avaluacions en curs i de rendiment, especialment per aquells algoritmes que canvien en el temps i desenvolupats en múltiples contextes¹¹¹⁴. Té dos objectius principals: requerir a les empreses que considerin des del primer moment l'impacte social dels sistemes algorítmics per treballar en mitigar-lo abans del desenvolupament, i crear documentació de les decisions i avaluacions que puguin donar suport a futures regulacions¹¹¹⁵.

En relació amb la discriminació, s'ha plantejat que podria emprar-se l'avaluació d'impacte algorítmica per detectar la discriminació i suprimir-la, no només en relació a col·lectius oprimits, sinó també per d'identificar patrons que serien impossibles de trobar únicament arran de la informació individual¹¹¹⁶. A Canadà, es requereixen auditories que incloguin problemàtiques de discriminació per part de les institucions públiques des de l'1 d'abril de 2020 i el Govern federal ha establert una plataforma d'avaluacions d'impacte algorítmiques¹¹¹⁷ per ajudar a les administracions¹¹¹⁸ -han d'omplir un qüestionari que determina el nivell d'impacte d'un sistema de decisió automatitzada que està dissenyat per ajudar departaments i agències a entendre millor i gestionar els riscos associats a aquests sistemes. El Govern dels Països Baixos promou les avaluacions d'impacte dels drets fonamentals i els algoritmes (en anglès Fundamental Rights and

¹¹¹³ SELBST, A. D., "An institutional view of Algorithmic Impact Assessments", *op. cit.*, 2021, pàg. 127.

¹¹¹⁴ KAMINSKI, M. E. i MALGIERI, G., "Algorithmic impact assessments...", *op. cit.*, 2021, pàg. 139.

¹¹¹⁵ SELBST, A. D., "An institutional view of Algorithmic Impact Assessments", *op. cit.*, 2021, pàg. 118.

¹¹¹⁶ KAMINSKI, M. E. i MALGIERI, G., "Algorithmic impact assessments...", *op. cit.*, 2021, pàg. 139.

¹¹¹⁷ GOVERNMENT OF CANADA, Algorithmic Impact Assessment Tool. Disponible a: <https://www.canada.ca/en/government/system/digital-government/digital-government-innovations/responsible-use-ai/algorithmic-impact-assessment.html>. Darrera visita el 2 de setembre de 2023.

¹¹¹⁸ DÉFENSEUR DES DROITS (DDD) ET COMMISSION NATIONALE INFORMATIQUE ET LIBERTÉS (CNIL), Algorithms: Preventing automation of discrimination, Paris, 2020, pàg. 10.

Algorithms Impact Assessment) que preveuen específicament les problemàtiques dels algoritmes per respectar el dret a la no discriminació¹¹¹⁹.

Les avaluacions d'impacte algorítmiques, a part de les avaluacions d'impacte de protecció de dades, tenen diverses precursoras com les avaluacions d'impacte del medi ambient (en anglès Environmental Impact Statements)¹¹²⁰ -que va instaurar la National Environmental Policy Act (NEPA) dels Estats Units d'Amèrica del Nord-; les avaluacions d'impacte dels drets humans (Human Rights Impact Assessments)¹¹²¹; les avaluacions d'impacte de la privacitat (Privacy Impact Assessments)¹¹²²; les avaluacions d'impacte ètiques (Ethical Impact Assessments)¹¹²³; les avaluacions d'impacte de vigilància (Surveillance Impact Assessments)¹¹²⁴; i les avaluacions d'impacte racials (Racial Impact Statements)¹¹²⁵.

Les propostes d'avaluacions d'impacte algorítmiques encaixen en tres categories: models basats en la NEPA, models basats en les avaluacions d'impacte de protecció de dades de l'RGPD o models basats en un qüestionari (model emprat al Canadà). El model NEPA implica una avaluació d'impacte molt detallada, potencialment d'un centenar de pàgines, que demana respostes precises a preguntes obertes per explicar el procés de disseny. Altres característiques són la transparència i la participació pública en un marc de notificació i observacions. El model de l'RGPD només requereix avaluar sistemàticament el programa i mostrar totes les mesures previstes per mitigar els riscos. Kaminski i Malgieri detallen que Eslovènia ja inclou en els requisits de l'article 22 de l'RGPD la inclusió d'avaluacions d'impacte algorítmiques¹¹²⁶. La major diferència amb el model NEPA és

¹¹¹⁹ GOVERNMENT OF THE NETHERLANDS, Fundamental Rights and Algorithms Impact Assessment (FRAIA). Disponible a: <https://www.government.nl/documents/reports/2021/07/31/impact-assessment-fundamental-rights-and-algorithms> Darrera visita el 2 de setembre de 2023.

¹¹²⁰Vegeu, per exemple: FROMKIN, A.M., "Regulating Mass Surveillance as Privacy Pollution: Learning from Environmental Impact Statements", *University of Illinois Law Review*, vol. 2015, núm. 5, 2015, pàg. 1713-1790.

¹¹²¹Vegeu, per exemple: ROARK, M.L., "Human Impact Statements", *Washburn Law Journal*, vol. 54, núm. 3, 2015, pàg. 649-682; JANSSEN, H. L., "Detecting New Approaches for a Fundamental Rights Impact Assessment to Automated Decision-Making", *International Data Privacy Law*, vol. 10, núm. 1, 2020, pàg. 76-106.

¹¹²²Vegeu, per exemple: CLARKE, R., "Privacy Impact Assessment: Its Origins and Development", *Computer Law and Security Review*, vol. 25, núm. 2, 2009, pàg. 123-135.

¹¹²³Vegeu, per exemple: WRIGHT, D. i FRIEDEWALD, M., "Integrating Privacy and Ethical Impact Assessments", *Science and Public Policy*, vol. 40, núm. 6, 2013, pàg. 755-766.

¹¹²⁴Vegeu, per exemple: WRIGHT, D. i RAAB, C.D., "Constructing a Surveillance Impact Assessment", *Computer Law and Security Review*, vol. 28, núm. 6, 2012, pàg. 613-626.

¹¹²⁵Vegeu, per exemple: ERICKSON, J., "Racial Impact Statements: Considering the Consequences of Racial Disproportionalities in the Criminal Justice System", *Washington Law Review*, vol. 89, núm. 4, 2014, pàg. 1425-1465.

¹¹²⁶KAMINSKI, M. E. i MALGIERI, G., "Algorithmic impact assessments...", *op. cit.*, 2021, pàg. 129-133.

la transparència, ja que malgrat la guia oficial de les avaluacions d'impacte de protecció de dades recomana fer públic un resum de l'avaluació, no és una obligació¹¹²⁷. El tercer model és el del govern de Canadà, que en la seva directiva relativa a les decisions automatitzades, especifica que les agències del govern que emprin la presa de decisions mitjançant sistemes algorítmics han de completar una avaluació d'impacte algorítmica abans d'implementar-la. L'avaluació d'impacte algorítmica consisteix en diverses qüestions sobre els riscos de les decisions, el sector i el grau d'explicació, entre altres¹¹²⁸. Aquest últim model està fet per ser emplenat abans i després de la implementació del sistema¹¹²⁹.

Hi ha diversos/es autors/es que plantegen avaluacions d'impacte algorítmiques amb les seves particularitats. Selbst proposa un ús de les avaluacions d'impacte algorítmiques per aplicar als departaments de policia que usen tecnologies de policia predictiva. S'haurien d'usar abans d'aplicar aquesta tecnologia havent d'analitzar diferents alternatives que el departament de policia hauria d'avaluar (s'hauria d'incloure una alternativa de no-acció), escollint l'alternativa preferida amb proposta de mesures de mitigació¹¹³⁰, fent èmfasi en la importància de revelar informació a la població i la supervisió judicial no només del procés sinó també del contingut.

Katyal incorpora elements de Selbst suggerint unes avaluacions d'impacte humanes en la presa de decisions algorítmica. Recomana un compromís amb la rendició de comptes algorítmica i l'antidiscriminació en relació amb el contingut i no només processal. Afegeix que les companyies haurien d'identificar poblacions afectades, l'efecte de la incertesa o error en aquests grups i l'estudi de si la decisió tindrà un efecte advers en un grup de població en particular¹¹³¹. Estableix diversos criteris que creu que s'haurien d'examinar en una avaluació d'impacte abans d'implementar un sistema algorítmic en base a diversos articles doctrinals¹¹³². Aquest criteris són:

¹¹²⁷ SELBST, A. D., "An institutional view of Algorithmic Impact Assessments", *op. cit.*, 2021, pàg. 140-143.

¹¹²⁸ *Ídem, ibidem.*

¹¹²⁹ *Ídem*, pàg. 147.

¹¹³⁰ *Ídem*, pàg. 173-178.

¹¹³¹ KATYAL, S. K., "Private Accountability in the Age of Artificial Intelligence", *op. cit.*, 2019, pàg. 115 i 116.

¹¹³² Vegeu: SELBST, A. D., "Disparate Impact in Big Data Policing", *op. cit.*, 2017, pàg. 109-195; ROARK, M.L., "Human Impact Statements", *op. cit.*, 2015, pàg. 649-682; DIAKOPOULOS, N. *et al.*, "Principles for Accountable Algorithms and a Social Impact Statement for Algorithms", 2016. Disponible a: <http://sorelle.friedler.net/papers/principles.pdf>. Darrera visita el 2 de setembre de 2023.

- a) identificar les poblacions potencialment afectades i determinar la seva raça, ètnia, sexe, orientació sexual, nació o altres categories;
- b) identificar l'efecte de la incertesa o error estadístic en diferents grups;
- c) estudiar si la decisió tindrà un efecte advers en el grup de població;
- d) explorar si hi ha alternatives raonables, menys discriminatòries o maneres de mitigar el dany, incloent la consideració de noves variables o altres dades;
- e) emprar altres tècniques de tractament de dades i nous mètodes d'avaluació;
- f) considerar cada alternativa en detall per tal que les persones avaluadores puguin valorar els mèrits comparatius;
- g) identificar i explicar l'alternativa preferida de l'entitat entre diverses opcions de disseny algorítmic¹¹³³.

L'AI Now Institute també presenta un informe on defensa la realització d'una avaluació d'impacte algorítmica abans que una agència pública utilitzi un sistema de decisió automatitzat. L'avaluació d'impacte algorítmica presentada requereix un període per revelar informació a la població i que es puguin fer observacions -admetent com a repte la regulació de la propietat industrial i intel·lectual¹¹³⁴-. També es planteja que s'hagi de renovar periòdicament i que s'hagi de permetre a persones investigadores i auditors la revisió del sistema una vegada estigui en funcionament¹¹³⁵.

En relació amb l'RGPD, Mantelero exposa la idea d'una avaluació d'impacte de drets humans, social i ètica (en anglès Human Rights, Social and Ethical Impact Assessments -HRSEIA-) en el context de la intel·ligència artificial. Un híbrid entre una avaluació d'impacte de Drets Humans (HRIA) i una avaluació d'impacte de privacitat (PIA), ja que observa que a la pràctica les avaluacions d'impacte de protecció de dades, inclús en el context europeu, han tendit a centrar-se en la qualitat de les dades i la seguretat de les

¹¹³³ KATYAL, S. K., "Private Accountability in the Age of Artificial Intelligence", *op. cit.*, 2019, pàg. 116 i 117.

¹¹³⁴ REISMAN, D., SCHULTZ, J., CRAWFORD, K. i WHITTAKER, M., *Algorithmic Impact Assessments: a Practical Framework...*, *op. cit.*, 2018, pàg. 16.

¹¹³⁵ *Ídem*, pàg. 8-16.

dades, deixant fora qualsevol impacte social i legal més ampli¹¹³⁶, i les avaluacions d'impacte socials o ètiques cobreixen una ampla gamma de temes, àmplies categories teòriques i solucions heterogènies¹¹³⁷. Suggereix un sistema basat en dos elements: un qüestionari d'autoavaluació i un comitè d'experts *ad hoc*, on les empreses voluntàriament tindrien present l'impacte ètic i social a més dels drets humans. Comenta la rellevància del rol d'una avaluació d'impacte d'aquestes característiques per adreçar les dimensions col·lectives dels perjudicis de les dades. S'encoratja, com en les altres propostes, a revelar informació a la població¹¹³⁸.

Villani planteja afegir les avaluacions d'impacte de discriminació (Discrimination Impact Assessments) a les avaluacions d'impacte que es fan en el marc de l'RGPD, per guiar les persones que dissenyen la intel·ligència artificial a interrogar-se sobre les conseqüències socials dels algoritmes que produeixen¹¹³⁹.

Ivanova defensa que l'avaluació d'un risc específic de discriminació hauria de ser un pas obligatori en el disseny i ús d'un sistema d'intel·ligència artificial (tant únicament automatitzat com parcial). El marc inclouria tres passos:

- 1) Una avaluació de la necessitat i proporcionalitat del tractament,
- 2) Una avaluació dels riscos pel dret a la no discriminació, i
- 3) L'aplicació de les mesures previstes per adreçar el risc¹¹⁴⁰.

Proposa un marc metodològic per avaluar els riscos del tractament automatitzat de dades pel dret a la no discriminació, donant a les persones responsables del tractament una guia indicativa sobre les principals fonts de risc i mesures potencials per prevenir i mitigar la discriminació algorítmica. Considera que l'avaluació d'impacte de protecció de dades podria esdevenir una avaluació d'impacte algorítmica salvaguardant més efectivament la privacitat de grup, la justícia substantiva i els valors ètics i interessos del conjunt de la societat¹¹⁴¹.

¹¹³⁶MANTELERO, A., "AI and Big Data: A Blueprint for a Human Rights, Social and Ethical Impact Assessment", *Computer Law and Security Review*, vol. 34, núm. 4, 2018, pàg. 761.

¹¹³⁷ *Idem*, pàg. 771.

¹¹³⁸ *Idem*, pàg. 769.

¹¹³⁹ VILLANI, C., *Donner un sens à l'intelligence artificielle...*, *op. cit.*, 2018, pàg. 147 i 148.

¹¹⁴⁰ IVANOVA, Y., "The Data Protection Impact Assessment...", *op. cit.*, 2020, pàg. 3 i 4.

¹¹⁴¹ *Idem*, pàg. 17 i 18.

Tant la proposta de Selbst com la de l'AI Now Institute requereixen que les avaluacions d'impacte algorítmiques siguin obligatòries i aborden únicament la seva aplicació en el sector públic; en canvi, la proposta de Katyal és que les avaluacions siguin voluntàries. Només Katyal considera les avaluacions d'impacte algorítmiques com una eina d'un sistema més gran de governança que interactua amb d'altres¹¹⁴². Malgrat que totes les propostes divergeixen en certs aspectes, presenten certes limitacions comunes.

4.4.2. Limitacions de les propostes d'avaluacions d'impacte algorítmiques

Una de les limitacions principals, que també hem esmentat en les avaluacions d'impacte de protecció de dades, la trobem si no s'obliga a compartir amb el públic general els resultats de les avaluacions d'impacte algorítmiques. Si no es fa aquest pas, es falla a activar terceres parts necessàries com actors de la societat civil¹¹⁴³. Involucrar de manera significativa les comunitats afectades¹¹⁴⁴ sembla plantejar-se com una tasca difícil. Mentre que moltes de les propostes d'avaluacions d'impacte algorítmiques preveuen la implicació de les comunitats afectades, cap de les propostes detalla de quina manera dur-ho a terme a la pràctica o anar més enllà d'una implicació a través d'observacions. El model NEPA confia en la notificació i les observacions però una oportunitat de comentar no permet resoldre els problemes que necessitarien una conversa¹¹⁴⁵.

S'ha de puntualitzar també una altra limitació, i és que qualsevol implementació de les avaluacions d'impacte algorítmiques ha d'aprendre del propi camp tècnic¹¹⁴⁶ (entendre com la intel·ligència artificial funciona a la pràctica, tal com està implementada per les empreses regulades, és crucial per entendre les limitacions del règim d'avaluacions d'impacte algorítmiques i fer-lo el més eficient possible). Selbst argumenta que els

¹¹⁴² KATYAL, S. K., "Private Accountability in the Age of Artificial Intelligence", *op. cit.*, 2019, pàg. 117.

¹¹⁴³ KAMINSKI, M. E. i MALGIERI, G., "Algorithmic impact assessments...", *op. cit.*, 2021, pàg. 134.

¹¹⁴⁴ En relació amb la inclusió de les comunitats afectades en la governança de la tecnologia, vegeu, per exemple: CITRON, D. K., "Technological Due Process", *op. cit.*, 2008, pàg. 1312 (l'autora proposa juntes de revisió de la tecnologia de la informació que ofereixin oportunitats a parts interessades i el públic en general per comentar sobre el disseny i entrenament d'un sistema); KATELL, M., YOUNG, M., DAILEY, D., HERMAN, B., GUETLER, V., TAM, A., BINZ, C., RAZ, D., i KRAFFT, P., "Toward situated interventions for algorithmic equity: Lessons from the field", *Proceedings of the 2020 Conference on Fairness, Accountability, and Transparency*, 2020, pàg. 45.

¹¹⁴⁵ SELBST, A. D., "An institutional view of Algorithmic Impact Assessments", *op. cit.*, 2021, pàg. 160.

¹¹⁴⁶ *Ídem*, pàg. 176.

objectius de les avaluacions d'impacte algorítmiques s'acompliran millor en col·laboració entre qui regula i la indústria que no des d'una perspectiva contradictòria¹¹⁴⁷, per tant s'ha de confiar inherentment en la cooperació amb el sector privat. Tanmateix, apunta, companyies tecnològiques han promès repetidament centrar les preocupacions ètiques i de drets humans entre les seves operacions i repetidament han fallat a aportar detalls sobre la seva implementació¹¹⁴⁸.

Per tant, hi ha riscos significatius per a l'eficàcia de la regulació, que en dependre d'un sistema de col·laboració, es veurà afectada per la cultura institucional i les prioritats de les empreses tecnològiques¹¹⁴⁹. A més, les avaluacions d'impacte algorítmiques han de ser ambigües en les seves definicions ja que aquest enfocament es defineix parcialment per la seva flexibilitat i la seva confiança en l'expertesa de l'empresa¹¹⁵⁰, fet que dificulta la seva regulació.

No obstant això, ens sembla interessant estudiar, finalment, la proposta que s'ha fet per part de la doctrina en relació a que aquestes avaluacions puguin fer-se servir per proporcionar documentació sobre l'impacte dels sistemes algorítmics en col·lectius oprimits dels sistemes algorítmics.

4.4.3. Consideració de les avaluacions d'impacte algorítmiques com a generadores de documentació sobre l'impacte dels sistemes algorítmics en determinats grups i col·lectius

Es proposa tant per part de Kaminski i Magliari¹¹⁵¹ com per part de Selbst¹¹⁵² que les avaluacions d'impacte algorítmiques serveixin com a connexió amb l'exercici de drets individuals, és a dir, que puguin servir com a generadores de documentació relativa a

¹¹⁴⁷ *Ídem*, pàg. 175.

¹¹⁴⁸ *Ídem*, pàg. 190.

¹¹⁴⁹ *Ídem*, pàg. 152.

¹¹⁵⁰ *Ídem*, pàg. 165.

¹¹⁵¹ KAMINSKI, M. E. i MALGIERI, G., "Algorithmic impact assessments...", *op. cit.*, 2021, pàg. 125-144.

¹¹⁵² SELBST, A. D., "An institutional view of Algorithmic Impact Assessments", *op. cit.*, 2021, pàg. 117-191.

l'impacte del sistema en col·lectius oprimits -elaborant la informació necessària per a una persona que s'ha vist discriminada per establir un cas *prima facie*.

Mentre el dret a una explicació individual en relació amb una decisió automatitzada pot esdevenir útil en algunes situacions, molts sistemes requereixen una anàlisi grupal¹¹⁵³. L'avaluació d'impacte algorítmica és una perspectiva sistèmica que requereix una anàlisi sobre com el sistema impacta no només sobre una persona, sinó sobre grups de persones¹¹⁵⁴. Per poder exercir els drets individuals és necessari que existeixi documentació -la mateixa que constituiria una avaluació d'impacte algorítmica- descrivint, per exemple, poblacions o situacions en les quals és menys probable que el sistema algorítmic funcioni de manera òptima¹¹⁵⁵. Selbst planteja que el requeriment de documentació *ex ante* serveixi com a possible remei *ex post*, és a dir, que les avaluacions d'impacte algorítmiques proveeixin la informació necessària per tal que la normativa antidiscriminació funcioni millor¹¹⁵⁶.

En el mateix sentit, i més específicament en el context de l'RGPD, Kaminski i Malgieri defensen que el contingut de l'avaluació d'impacte de protecció de dades sota l'RGPD no dista massa dels drets individuals de transparència en relació amb les decisions preses mitjançant algoritmes. Per exemple, quan l'article 35.7 requereix que l'avaluació d'impacte de protecció de dades contingui una descripció sistemàtica de les operacions de tractament previstes es podria utilitzar per donar a les persones informació significativa de la lògica del sistema. Hi ha nombroses similituds entre la informació que una persona responsable de dades ha de generar durant una avaluació d'impacte i la informació que ha de revelar com a part dels deures de transparència algorítmica de l'RGPD. La tasca de descriure les operacions de tractament és similar a la tasca de clarificar les categories de dades personals i com es construeix l'elaboració de perfils. La tasca d'avaluar els riscos i efectes en les persones és similar al deure d'explicar l'impacte de l'ús de l'elaboració de perfils. Finalment, el deure d'establir mesures dels drets individuals en el cas de les decisions preses mitjançant algoritmes és similar al deure de descriure les mesures previstes per minimitzar els riscos en les avaluacions d'impacte¹¹⁵⁷.

¹¹⁵³ REISMAN, D., SCHULTZ, J., CRAWFORD, K. i WHITTAKER, M., *Algorithmic Impact Assessments: a Practical Framework...*, op. cit., 2018, pàg. 19.

¹¹⁵⁴ KAMINSKI, M. E. i MALGIERI, G., "Algorithmic impact assessments...", op. cit., 2021, pàg. 138.

¹¹⁵⁵ SELBST, A. D., "An institutional view of Algorithmic Impact Assessments", op. cit., 2021, pàg. 137.

¹¹⁵⁶ *Ídem*, pàg. 131.

¹¹⁵⁷ KAMINSKI, M. E. i MALGIERI, G., "Algorithmic impact assessments...", op. cit., 2021, pàg. 140.

També, combinar les tasques de les avaluacions d'impacte i la transparència individual pot ajudar a les persones responsables del tractament, atès que optimitzarien els esforços que -sense fer aquest exercici- s'emprarien en dues tasques diferents. Publicant els resultats (o un resum) de les avaluacions d'impacte, per una banda, durien a terme una bona pràctica i es podria considerar complert el deure de buscar el punt de vista de les persones afectades (article 35.9 RGPD), i per l'altra banda, durien a terme una mesura adequada per protegir drets i deures de les persones (article 22.3 i article 35.7.d de l'RGPD) enriquint la transparència amb mesures de rendició de comptes¹¹⁵⁸.

En definitiva, els mecanismes de responsabilitat proactiva, sota certes condicions, poden ser efectius per protegir les persones de la discriminació algorítmica i donar-los eines per rebatre-les. La implementació d'una política de compliment que requereixi una monitorització constant del sistema algorítmic i aclareixi la responsabilitat de l'acció discriminatòria, l'obligació de prendre mesures per disminuir la discriminació que puguin causar els sistemes amb codis de conducta, certificacions, segells i marques que es supervisin periòdicament, la perspectiva de protecció des del disseny i per defecte, i la realització d'avaluacions d'impacte algorítmiques -derivades de les avaluacions d'impacte de protecció de dades- que puguin, a part de detectar col·lectius que es poden veure perjudicats pel sistema algorítmic, produir documentació que sigui efectiva i necessària per l'establiment d'un cas *prima facie* de discriminació, són aportacions que l'RGPD pot plantejar per la detecció i mitigació de la discriminació algorítmica. No obstant això, hi ha grans obstacles a salvar, com el cost d'una supervisió òptima d'un sistema que pot evolucionar amb el temps, la no obligació de revelar al públic general el resultat de les auditories o avaluacions d'impacte i la necessitat de confiar en el sector privat per tal que la regulació funcioni.

A continuació, estudiem si l'article 80 de l'RGPD, que dona la possibilitat d'una defensa dels drets col·lectius en referència a una vulneració de les obligacions en matèria de protecció de dades personals, pot millorar la protecció envers la discriminació algorítmica, tenint en compte les limitacions ja esmentades relatives a la protecció individual.

¹¹⁵⁸ *Ídem*, pàg. 143 i 144.

5. Vessant col·lectiva de la protecció envers la discriminació: Article 80 RGPD

Com hem esmentat en l'apartat cinquè del capítol III, la protecció individual sobre la qual s'erigeix l'estructura de protecció de les persones de l'RGPD té certes limitacions per ser efectiva en relació amb la discriminació algorítmica. En aquest apartat, analitzem les possibilitats que ofereix l'article 80.2 de l'RGPD relatiu a la defensa dels drets de la persona per part d'una entitat, organització o associació, arran de la interpretació feta pel TJUE en la sentència de 28 d'abril de 2022, *Meta Platforms Ireland Limited*¹¹⁵⁹ (5.1), que possibilita la defensa de drets des de la perspectiva col·lectiva. Aquesta sentència aclareix els mecanismes per a una representació col·lectiva de la protecció envers la discriminació (5.2), però no es poden obviar els obstacles que alhora s'erigeixen per l'exercici efectiu dels drets conferits en l'article (5.3).

5.1. Article 80.2 RGPD: Representació sense mandat exprés

L'article 80 de l'RGPD, en el primer apartat, reconeix el dret de la persona interessada a encomanar a una entitat, organització o associació -que compleixi amb els requisits estipulats- que en el seu nom presenti una reclamació davant d'una autoritat de control o òrgan jurisdiccional i exerceixi el seu dret a ser indemnitzada (article 82 RGPD) i, en el segon apartat, estableix que qualsevol estat membre podrà disposar que qualsevol entitat, organització o associació, amb independència del mandat de la persona interessada, tingui dret a presentar una reclamació o a exercir el seu dret a la tutela judicial efectiva en el mateix estat membre davant l'autoritat de control competent¹¹⁶⁰. S'ha de puntualitzar que el poder legislatiu espanyol no ha regulat aquesta matèria explícitament¹¹⁶¹ ja que l'article

¹¹⁵⁹ TJUE (Sala Tercera), sentència de 28 d'abril de 2022, C-319/20, *Meta Platforms Ireland Limited*, anteriorment *Facebook Ireland Limited*, contra *Bundesverband der Verbraucherzentralen und Verbraucherverbände - Verbraucherzentrale Bundesverband e.V.*

¹¹⁶⁰ Article 80 de l'RGPD. En la doctrina espanyola, vegeu: VILLALBA CANO, L., "La representació de los interesados (Comentario al artículo 80 RGPD)", a TRONCOSO REIGADA, A., (dir.) i VILLALBA CANO, L., (coord.), *Comentario al Reglamento General de Protección de Datos y a la Ley Orgánica de Protección de Datos Personales y Garantía de los Derechos Digitales*, Tomo II, Pamplona, Aranzadi, 2021, pàg. 3019-3040.

¹¹⁶¹ PALMA ORTIGOSA, A. *Decisiones automatizadas y protección de datos. Especial atención a los sistemas de inteligencia artificial*, Madrid, Dykinson, 2022, pàg. 326.

80 no ha estat objecte de desenvolupament en la LOPDGDD, i per tant, ens hem d'adreçar a la regulació general que trobem a Ley de Enjuiciamiento Civil (d'ara en endavant, LEC)¹¹⁶².

Aquest article implica que un dret de caràcter personal, com és el dret a la protecció de dades personals, l'exercici del qual correspon, en general, a la persona interessada, podrà ser defensat per una entitat, associació o organització a la qual es confereixi aquest mandat¹¹⁶³, facilitant així a les persones físiques l'exercici dels seu dret a la tutela judicial efectiva de la mateixa manera que no s'impedeix ni dificulta l'exercici dels drets de manera individual. En referència a si les entitats, associacions o organitzacions poden actuar sense el mandat individual de les persones afectades¹¹⁶⁴, van sorgir dubtes per saber l'abast de la representació estipulada en l'article 80.2.

Aquests dubtes van ser aclarits en la sentència del TJUE, de 28 d'abril de 2022, Meta Platforms Ireland Limited. Meta Platforms Ireland Limited posava a disposició de les persones usuàries jocs gratuïts subministrats per tercers, i mostrava a la persona usuària l'avertència que l'ús de l'aplicació permetia a la societat que facilitava els jocs obtenir determinades dades personals i, l'autoritzava a realitzar publicacions en nom seu¹¹⁶⁵, versant la qüestió prejudicial sobre si una entitat de consumidors podia actuar contra l'infractor sense que hi hagués un mandat d'una persona interessada. El Tribunal va avalar que les associacions de consumidors exercitessin accions, sense mandat i amb independència de la vulneració de drets concrets de la persona interessada, contra qui infringís la normativa en matèria de protecció de dades¹¹⁶⁶.

El Tribunal va considerar que no calia mandat concret, atès que a l'entitat que exercia l'acció de representació (que també anomenem en el treball acció col·lectiva) no podia

¹¹⁶²Article 11 i article 11 bis de la LEC.

¹¹⁶³HERRÁN ORTIZ, A. I., "Aproximación al derecho a la protección de datos personales en Europa. El Reglamento General de Protección de Datos personales a debate", *Revista de Derecho, Empresa y Sociedad*, núm. 8, 2016, pàg. 182.

¹¹⁶⁴La possibilitat d'indemnització està contemplada en l'article 82 de l'RGPD. El considerant 142 de l'RGPD especifica que la persona interessada haurà de tenir el dret a conferir mandat a una entitat, organització o associació per presentar en el seu nom una reclamació davant l'autoritat de control, per exercir el seu dret a la tutela judicial efectiva o, si així ho estableix l'estat membre, per exercir el dret a rebre una indemnització en el seu nom. Posteriorment exposa que l'entitat, organització o associació no pot estar autoritzada a reclamar una indemnització en nom d'una persona interessada al marge del mandat d'aquesta.

¹¹⁶⁵ TJUE (Sala Tercera), sentència de 28 d'abril de 2022, C-319/20, Meta Platforms Ireland Limited, anteriorment Facebook Ireland Limited, contra Bundesverband der Verbraucherzentralen und Verbraucherverbände - Verbraucherzentrale Bundesverband e.V, par. 34.

¹¹⁶⁶ *Ídem*, par. 83.

exigir-se-li que prèviament dugués a terme una identificació individual de la persona que tingué específicament la condició de persona interessada per un tractament de dades suposadament contrari a l'RGPD. El TJUE apuntava que el concepte "interessat/da" segons l'article 4.1 de l'RGPD compren no només una "persona física identificada" sinó que també una "persona física identificable", és a dir, una persona la identitat de la qual pugui determinar-se directament o indirectament, per exemple, amb un nom, un número d'identificació, dades de localització o un indicador en línia. El tribunal raonava llavors, que la designació d'una categoria o grup de persones interessades pel tractament també podia ser suficient a efectes de l'acció de representació¹¹⁶⁷.

A més, va determinar que l'exercici d'aquesta acció no estava supeditat a l'existència d'una vulneració concreta dels drets que s'atorguen a les persones interessades, sinó que l'acció només pressuposava que l'entitat considerava que els drets de les persones interessades havien estat vulnerats pel tractament de dades i per tant, era contrari a les disposicions de l'RGPD. Per consegüent, va concloure que en virtut de l'article 80.2 només calia al·legar que el tractament en qüestió podia afectar els drets de persones físiques identificades o identificables, sense que fos necessari provar un perjudici real patit a causa de la vulneració d'aquests drets¹¹⁶⁸, i que aquest article no s'oposava a una normativa nacional que permetés a una associació en defensa dels interessos de les persones consumidores exercir accions judicials sense mandat i amb independència de la vulneració dels drets concrets de la persona interessada¹¹⁶⁹.

A Espanya, aquesta possibilitat existeix atès que l'apartat quart de l'article 15 de la LEC¹¹⁷⁰ estableix que per exercir-se una acció de cessació en defensa d'interessos col·lectius i dels interessos difusos de persones consumidores i usuàries, no és necessària la comunicació prèvia a les persones interessades de la notícia que el procés s'inicia ni l'oferta d'intervenció litisconsorcial. Per tant, sense mandat exprés, una entitat, organització o associació espanyola pot dur a terme una acció de cessació d'una infracció de protecció de dades (que afecti el dret de les persones consumidores).

¹¹⁶⁷ *Ídem*, par. 67, 68 i 69.

¹¹⁶⁸ *Ídem*, par. 70, 71, 72.

¹¹⁶⁹ *Ídem*, par. 83.

¹¹⁷⁰ Article 15.4 de la LEC.

5.2. Possibilitats de l'article 80.2 RGPD per redreçar la discriminació algorítmica

La interpretació del TJUE respecte l'article 80.2 de l'RGPD és molt positiva en relació amb la rellevància que pot assolir en referència a la lluita envers la discriminació algorítmica. Suposa que malgrat no existir víctimes de discriminació identificades, el simple fet que un algoritme pugui ser discriminatori arran del tractament de dades, pot ocasionar que es qüestioni si vulnera la normativa de protecció de dades mitjançant la denúncia per part d'una entitat, organització o associació.

En el cas que un tractament de dades discriminatori no entrés en l'àmbit d'aplicació de l'article 22 en no tractar-se d'una decisió únicament automatitzada, els resultats generats per dades d'entrenament esbiaixades violen el principi d'adequació de les dades (art. 5.1.d de l'RGPD) i la discriminació algorítmica no justificada també infringeix l'equitat en el tractament de dades (art. 5.1.a de l'RGPD)¹¹⁷¹, per tant, la discriminació algorítmica pot suposar una vulneració de les obligacions establertes en l'RGPD, sense haver-se de basar exclusivament en l'article 22 de l'RGPD.

Com ja hem esmentat, la vessant individual de la protecció, quan ens referim a la discriminació algorítmica, podrà ser efectiva si ens trobem davant el cas de discriminació directa. Com hem esmentat, si s'ha discriminat en base a un perfil explícitament relacionat amb una categoria sensible i de manera evident, es podrà comptar amb que els drets individuals conferits en l'RGPD siguin efectius en si mateixos. És a dir, si el perfil és explícit, es podrà arribar a la conclusió, per exemple, a través dels drets individuals conferits en l'article 13.2 (f), 14.2 (g) i 15.1 (h) a conèixer la lògica del sistema (si la informació que es dona és comprensible), que s'han utilitzat criteris basats en dades sensibles¹¹⁷², i es podrà invertir llavors la càrrega de la prova en un cas de discriminació *prima facie*, ja que la informació sobre una única persona serà suficient per evidenciar que s'ha produït discriminació.

¹¹⁷¹ HACKER, P., "Teaching fairness to artificial intelligence...", *op. cit.*, 2018, pàg. 1168.

¹¹⁷² Part de la doctrina defensa que en base a un dret a l'explicació del funcionament de l'algoritme es podria provar la discriminació *prima facie*. Tot i així, tal com remarquem al llarg d'aquest treball, la persona que s'ha vist discriminada realment hauria de disposar d'evidències suficients relatives a que s'ha emprat un perfil explícitament basat en una categoria sensible, i això, a causa de l'opacitat i la manca de transparència d'aquests sistemes, pot ser difícil.

En canvi, en relació amb la discriminació indirecta, la vessant individual de la protecció serà insuficient, ja que es necessitaran dades sobre les tendències a la discriminació i els biaixos de l'algoritme -essent necessari disposar de dades de molts casos per poder valorar l'existència de discriminació- no inclosos en la informació d'accés individual (a la qual, també sigui dit, només es tindrà accés quan la decisió sigui únicament automatitzada), de manera que serà molt més difícil establir una discriminació *prima facie*.

Davant d'aquesta perspectiva, creiem veure en l'article 80.2 de l'RGPD, mitjançant una acció de cessació d'un tractament de dades que ocasioni discriminació duta a terme per una entitat, organització o associació, la possibilitat de suplir o complementar les mancances de la protecció individual.

L'article 80.2 arran de la interpretació feta per part del TJUE en la seva sentència de 28 d'abril de 2022 presenta un escenari molt favorable en relació amb la discriminació algorítmica. Aquesta acció, que podria provocar la cessació de l'ús d'un sistema algorítmic amb tendència a la discriminació i al biaix, podria ser molt rellevant en el cas que les entitats, organitzacions i associacions tinguessin informació grupal sobre l'impacte de cada determinat sistema algorítmic en certs grups i col·lectius de població. De la mateixa manera que s'ha fet amb l'assumpte referenciat de protecció de dades, en el moment en què fos pública certa tendència a discriminar (o inclús, certa infrarepresentació d'algun col·lectiu poblacional per part d'un sistema algorítmic), una entitat, organització o associació podria a través de l'article 80.2 de l'RGPD dur a terme una acció de cessació desvinculada de qualsevol vulneració concreta del dret a la protecció de dades d'una persona interessada sense la necessitat d'un mandat de tal persona. No obstant això, no volem obviar els obstacles amb que es poden trobar entitats, organitzacions i associacions per l'exercici de l'article 80.2 de l'RGPD.

5.3. Dificultats per l'exercici de l'article 80 en relació amb la discriminació algorítmica

En primer lloc, i tornant a la problemàtica de l'accés a la informació, és cert que identificar les condicions per a l'exercici d'aquests drets requereix informació que potser no és immediatament accessible a causa de les diverses formes d'opacitat que comporten els

sistemes de decisió automatitzada¹¹⁷³. Per dur a terme una acció col·lectiva s'entén que, o bé l'entitat haurà de fer una investigació per aconseguir prova de discriminació, o bé la persona interessada haurà de realitzar tant l'esforç inicial d'identificar una irregularitat en la presa de decisions com de localitzar i contactar amb l'associació perquè aquesta la representi¹¹⁷⁴.

A més, tot i que aquestes entitats, associacions o organitzacions puguin arribar a tenir un pes rellevant i un coneixement suficient en algun sector, no seran realment efectives sense capacitats tècniques per realitzar l'anàlisi de dades¹¹⁷⁵. La informació hauria d'estar disponible sense que les entitats, organitzacions o associacions haguessin d'emprar esforços tècnics per obtenir-la. És a dir, seran necessàries eines de transparència efectives per poder presentar proves de la presumpta discriminació. Per exemple, l'assumpte Meta Platforms Ireland Limited (on l'entitat havia localitzat clarament en quin sentit Meta Platforms Ireland Limited podia estar infringint la normativa de protecció de dades) actualment no sembla extrapolable a la discriminació algorítmica (on pot ser molt més complicat -o directament impossible- tenir aquesta informació). Per consegüent, en el moment en el que existís informació grupal accessible, l'article 80.2 de l'RGPD tal com està interpretat pel TJUE en la sentència analitzada, facilitaria molt la protecció de les persones susceptibles de ser discriminades, a través d'accions de cessació de tractament discriminatori.

En segon lloc, l'àmbit d'incidència del tractament massiu de dades no té límits territorials, contràriament als requisits per a la protecció dels drets en sentit col·lectiu, que sí que en tenen. Cada estat membre pot estipular requisits addicionals a aquells establerts en l'RGPD, com, per exemple períodes mínims d'activitat¹¹⁷⁶ o exigències en relació a l'estat membre on s'ha constituït l'entitat, associació o organització¹¹⁷⁷, causant disparitat de legislacions i dificultant l'ús d'aquest mètode de protecció enfront les discriminacions.

¹¹⁷³Vegeu, per exemple: BURRELL J., "How the machine 'thinks': Understanding opacity...", *op. cit.*, 2016, pàg. 1-12.

¹¹⁷⁴ROIG, A., *Las Garantías frente a las decisiones automatizadas...*, *op. cit.*, 2020, pàg. 127.

¹¹⁷⁵*Idem*, pàg. 130.

¹¹⁷⁶Per exemple, a França, l'article 37 (IV) de la Loi n° 78-17 du 6 janvier 1978 relative à l'informatique, aux fichiers et aux libertés, disposa que només podran exercir l'acció col·lectiva aquelles associacions inscrites regularment després d'almenys cinc anys que tinguin en el seu estatut la protecció de la vida privada o la protecció de les dades de caràcter personal.

¹¹⁷⁷Per exemple, en l'Estat belga, l'article 220 de la Loi du 30 juillet 2018 relative à la protection des personnes physiques à l'égard des traitements de données à caractère personnel s'exigeix que l'entitat, associació o organització estigui correctament constituïda conforme al dret belga.

Així mateix, cada estat membre pot divergir en l'ús de les accions col·lectives i el nombre d'associacions que s'hi dediquen, de manera que en alguns estats membres la protecció es pot donar més fàcilment que en d'altres. També, malgrat que a través de l'article 80.2, les entitats, organitzacions o associacions podran exercir una acció de cessació sense el mandat de les persones interessades, no serà així per exercir el dret a la indemnització a la que es fa referència en l'article 82 de l'RGPD¹¹⁷⁸, com especifica l'article 80.1¹¹⁷⁹, on sí que serà necessari el mandat de la persona interessada.

Finalment i en tercer lloc, la capacitat dels algorismes de classificar en base a grups de persones que comparteixen les mateixes característiques però que desconeixen que formen part d'un grup, sense ser conscients de la identitat d'altres membres del grup (és a dir, amb una percepció limitada de les seves problemàtiques col·lectives), pot dificultar que siguin conscients de la situació en la qual es troben col·lectivament. També pot ser difícil la seva efectivitat en el cas de sistemes algorítmics que directament exclouin les persones. En el cas que sí que fossin conscients de la seva pertinença al grup discriminat, aquesta seria la manera de poder fer efectiva la protecció envers la discriminació indirecta, ja que el fet de presentar diversos casos amb el mateix patró ocasionaria que es pogués determinar la discriminació *prima facie*.

A tall de recapitulació, la possibilitat que entitats, associacions i organitzacions puguin dur a terme accions de cessació en l'àmbit de la protecció de dades pot oferir eines per desafiar la discriminació algorítmica. Sobretot, després que el TJUE en la seva sentència de 28 d'abril de 2022, *Meta Platforms*, aclarís la possibilitat que la defensa es produeixi sense mandat i amb independència de la vulneració de drets concrets de la persona interessada. No obstant això, per tal que l'article 80.2 de l'RGPD pugui ser útil per desafiar la discriminació algorítmica calen eines de transparència efectives, salvar la dificultat que comporta que el grup de població que està essent discriminat no en sigui conscient, i que l'àmbit d'incidència del tractament massiu de dades no tingui límits territorials i en canvi, els requisits per a la protecció dels drets en sentit col·lectiu estiguin delimitats territorialment.

¹¹⁷⁸ Article 82 de l'RGPD.

¹¹⁷⁹ Article 80.1 de l'RGPD.

6. Conclusions parcials

En aquest capítol ens proposàvem respondre, en part, a la tercera qüestió de recerca del treball, que plantejava si algunes garanties de la regulació antidiscriminació i de protecció de dades podrien ser vàlides per protegir les persones de la discriminació algorítmica, donades certes condicions. Hem presentat els punts forts que presenta el marc regulador europeu de protecció de dades i si aquests, sota certes condicions, poden mitigar els punts febles que hem plantejat en el capítol III o directament ser útils per a l'abordatge de la discriminació algorítmica.

En primer lloc, la governança col·laborativa subjacent en l'RGPD pot ser un bon sistema per abordar les discriminacions algorítmiques, ja que inclou els recursos i l'expertesa del sector privat en la regulació pública. Nogensmenys, hi ha riscos molt rellevants i crucials per la protecció de drets fonamentals: que el sector privat es converteixi *de facto* en autoregulat, ja sigui perquè amb la seva expertesa i recursos pot definir i incidir discrecionalment en conceptes clau de la regulació, ja sigui perquè el poder públic no pot instaurar una rendició de comptes eficient.

En segon lloc, tant l'exigència d'exactitud com els articles 15, 16, 17, 18, 20 i 21 de l'RGPD, poden participar a millorar la qualitat de les dades, que tenen una incidència directa en els biaixos algorítmics causants de la discriminació algorítmica. No obstant això, l'RGPD planteja la qualitat de dades de manera estàtica, és a dir, les dades que hi ha es poden canviar, la qual cosa no s'adapta als sistemes algorítmics que evolucionen a través de la correlació de dades que no es podran alterar retrospectivament. A més, només genera eines per fer front a la vessant tècnica d'aquesta problemàtica, havent altres accions necessàries per mitigar l'aparició de biaixos i reduir-ne l'impacte, com la sensibilització social envers els biaixos inconscients, equips de treball amb diversitat d'edat, ètnia, cultura o gènere, ús de dades el més representatives possible, promoure IA diversa i inclusiva, comprendre i mesurar biaixos, generar noves bases de dades d'entrenament o que les persones que creen bases de dades en defineixin les característiques i les limitacions per tal que quan s'utilitzin es tingui tota la informació pertinent per al seu ús.

En tercer lloc, hem estudiat sota quines condicions la informació que proveeix l'RGPD podria ser útil en relació amb la discriminació algorítmica. El dret a l'explicació pot ser útil si es produeix discriminació directa (atès que amb la informació individual obtinguda la persona podria tenir suficients indicis com per presentar un cas *prima facie*), però si es produeix discriminació indirecta, malgrat que individualment pugui donar-se certa informació, sempre serà necessària una altra perspectiva grupal, basada en proves estadístiques o protecció col·lectiva. A més, hem argumentat per què la informació individual que es proporioni hauria d'estar estructurada com a explicació contrafàctica, segons el nostre parer, la tipologia més útil i senzilla per tal que una persona pugui detectar que s'ha vist discriminada. També hem analitzat les exigències de transparència que presenta l'RGPD, que podrien ser útils, per exemple, si s'empres per identificar sistemes algorítmics (és a dir, denominar i diferenciar cada sistema algorítmic de presa de decisió) i així, les persones que s'haguessin vist discriminades a través d'aquest, poguessin identificar-lo.

En quart lloc, hem estudiat com la política de responsabilitat proactiva de l'RGPD pot ajudar tant a aclarir qüestions de rellevància com a mitigar efectes de la discriminació algorítmica. Pot aclarir qüestions sobre la responsabilitat dels efectes discriminatoris del sistema algorítmic, que sota la nostra perspectiva han de recaure sobre l'empresa o organització que el posa en pràctica, llevat que la persona responsable de tractament els hagi causat per manca de diligència. Per mitigar els resultats discriminatoris de sistemes algorítmics que s'apliquen en contextos socials, la inclusió de criteris sobre la no discriminació en codis de conducta i com a requisits per obtenir una certificació, segell o marca pot ser una bona opció. Finalment, hi ha eines de responsabilitat proactiva de l'RGPD que han d'anar més enllà per poder ser útils en relació amb la discriminació algorítmica. És el cas de la protecció de dades des del disseny i per defecte, que hauria d'incorporar la igualtat des del disseny i per defecte com a estàndard en el desenvolupament dels algoritmes (per tal d'estructurar-los de manera que siguin sensibles a les formes predominants de discriminació i inclús corregeixin desequilibris socials entre membres de grups avantatjats i desavantatjats) i les avaluacions d'impacte de protecció de dades, que en el cas de veure's implicats algoritmes haurien de ser avaluacions d'impacte algorítmiques, que poguessin, a part de detectar col·lectius que es poden veure perjudicats pel sistema algorítmic, produir documentació que pugui ser efectiva i necessària per a l'establiment d'un cas *prima facie* de discriminació. No obstant això, hi

ha grans obstacles a salvar, com el cost d'una supervisió òptima d'un sistema que pot evolucionar amb el temps, la no obligació de revelar al públic general el resultat de les auditories o avaluacions d'impacte o la necessitat de confiar en el sector privat per tal que la regulació funcioni.

Finalment, en cinquè lloc, la possibilitat que ofereix l'article 80.2 de l'RGPD que entitats, associacions i organitzacions puguin dur a terme accions de cessació en l'àmbit de la protecció de dades pot oferir eines per desafiar la discriminació algorítmica, sobretot quan, havent-se habilitat a escala nacional, no sigui necessari el mandat ni la vulneració de drets concrets de la persona interessada. No obstant això, per tal que l'article 80.2 de l'RGPD pugui ser útil per desafiar la discriminació algorítmica calen eines de transparència efectives, salvar la dificultat que comporta que el grup de població que està essent discriminat en sigui conscient, i que l'àmbit d'incidència del tractament massiu de dades no tingui límits territorials i en canvi els requisits per la protecció dels drets en sentit col·lectiu estiguin delimitats territorialment.

Per tant, concloem, que malgrat la normativa de protecció de dades europea pugui mitigar, sota certes condicions, els punts febles que presenta per protegir les persones envers la discriminació algorítmica, i oferir possibilitats per iniciar litigis davant d'un tribunal, s'han de donar moltes circumstàncies que en l'actualitat no es donen ni sembla que s'hagin de donar (p. e. rebre informació útil per iniciar un plet de discriminació, que les persones tinguin un control real i efectiu de les seves dades o que siguin conscients que estan essent discriminades).

Si reprenem la tercera qüestió de recerca que ens plantejàvem en aquest treball i que hem abordat tant en aquest capítol com en l'anterior, a saber, si algunes garanties de la regulació antidiscriminació i de protecció de dades podrien ser vàlides per protegir les persones de la discriminació algorítmica, donades certes condicions, concloem que efectivament, sí que hi ha certs elements de les dues normatives que podrien ser vàlids per desafiar la discriminació algorítmica. Tanmateix, no considerem que constitueixin règims protectors de suficient entitat com per donar eines adequades a les persones que es vegin discriminades per desafiar les decisions dels sistemes algorítmics amb les quals s'ha causat la discriminació. En conseqüència, hem d'estudiar si la regulació de la intel·ligència artificial europea participa a conformar una normativa que ofereixi una millor protecció.

Capítol VI. Discriminació algorítmica i PRIA

Per acabar aquesta segona part del treball, en aquest capítol responem la quarta qüestió de recerca que ens plantejàvem, a saber, si el marc normatiu que planteja la UE per regular la intel·ligència artificial ofereix noves garanties per fer front a la discriminació algorítmica. Per fer-ho, estudiem la PRIA. Cal fer esment també, malgrat que no l'estudiem amb profunditat, a l'esborrany de la Convenció sobre intel·ligència artificial, drets humans, democràcia i estat de dret¹¹⁸⁰ del Consell d'Europa, que estableix certs principis, normes i drets per regular l'aplicació de sistemes d'intel·ligència artificial.

La PRIA va ser presentada per la Comissió Europea el 21 d'abril de 2021¹¹⁸¹. El 14 de juny de 2023 el Parlament Europeu va adoptar certes esmenes a la proposta de Reglament presentada per la Comissió¹¹⁸² que puntualitzarem, quan siguin rellevants, al llarg del capítol. La proposta de Reglament forma part d'un conjunt de propostes que s'han d'entendre conjuntament¹¹⁸³, i està orientada a harmonitzar la legislació en matèria d'intel·ligència artificial des d'una perspectiva de seguretat dels productes que s'introdueixen en el mercat intern de la UE (es persegueix una perspectiva horitzontal que, a diferència d'altres legislacions de seguretat dels productes, no depèn del sector,

¹¹⁸⁰ COMMITTEE ON ARTIFICIAL INTELLIGENCE (CAI), *Revised Zero Draft (Framework) convention on artificial intelligence, human rights, democracy and the rule of law*, CAI(2023)01, Council of Europe, Strasbourg, 6 January 2023.

¹¹⁸¹ COMISIÓN EUROPEA, *Propuesta de Reglamento del Parlamento Europeo y del Consejo relativo a las máquinas y sus partes y accesorios*, COM(2021) 202 final, Bruselas, 21 de abril de 2021.

¹¹⁸² EUROPEAN PARLIAMENT, *Amendments adopted by the European Parliament on 14 June 2023 on the proposal for a regulation of the European Parliament and of the Council on laying down harmonised rules on artificial intelligence (Artificial Intelligence Act) and amending certain Union legislative acts (COM(2021)0206 – C9-0146/2021 – 2021/0106(COD))1*, P9_TA(2023)0236, Brussels, 14 June 2023.

¹¹⁸³ Reglamento (UE) 2022/2065 del Parlamento Europeo y del Consejo de 19 de octubre de 2022 relativo a un mercado único de servicios digitales y por el que se modifica la Directiva 2000/31/CE (Reglamento de Servicios Digitales). DOUE, L 277/1, 27.10.2022; Reglamento (UE) 2022/1925 del Parlamento Europeo y del Consejo, de 14 de septiembre de 2022, sobre mercados disputables y equitativos en el sector digital y por el que se modifican las Directivas (UE) 2019/1937 y (UE) 2020/1828 (Reglamento de Mercados Digitales). DOUE, L 265/1, 12.10.2022; Reglamento (UE) 2022/868 del Parlamento Europeo y del Consejo de 30 de mayo de 2022 relativo a la gobernanza europea de datos y por el que se modifica el Reglamento (UE) 2018/1724 (Reglamento de Gobernanza de Datos). DOUE, L152/1, 3.6.2022; COMISIÓN EUROPEA, *Propuesta de Directiva del Parlamento Europeo y del Consejo, relativa a la responsabilidad por productos defectuosos*, COM(2022) 495 final, Bruselas, 28 de septiembre de 2022; COMISIÓN EUROPEA, *Propuesta de Directiva del Parlamento Europeo y del Consejo, relativa a la adaptación de las normas de responsabilidad no contractual a la inteligencia artificial*, COM (2022) 496 final, Bruselas, 28 de septiembre de 2022.

sinó que es basa en l'ús de la IA en general¹¹⁸⁴).

Es parteix d'una classificació dels sistemes d'IA segons els riscos que poden causar a la sanitat, la seguretat i els drets fonamentals. Aquest enfocament, planteja una visió més àmplia d'aquella que oferia la normativa de protecció de dades en la UE. En el text, la intel·ligència artificial es caracteritza com una oportunitat a la qual no es pot renunciar pels beneficis associats a la seva operativitat i eficiència, però que pot vulnerar o posar en risc els drets fonamentals i béns jurídics protegits¹¹⁸⁵. En aquest sentit, la UE afirma que no només focalitza la seva atenció en la innovació i el creixement econòmic, sinó també en les implicacions socials i ètiques, apuntant que el compliment dels principis ètics europeus, els requisits legals i els valors socials són essencials per crear un ecosistema de confiança¹¹⁸⁶.

Es defineix un sistema d'IA com un software que es desenvolupa mitjançant estratègies d'aprenentatge automàtic basades en la lògica i el coneixement, o estadístiques que poden, per un conjunt determinat d'objectius definits per éssers humans, generar resultats com contingut, prediccions, recomanacions o decisions que influeixen els ambients amb els quals interactua¹¹⁸⁷. Es fa la intersecció d'una definició funcional amb una definició descriptiva. L'amplitud de la definició potser no abasta tots els softwares però en captura la majoria. El Reglament és aplicable a proveïdors¹¹⁸⁸ i usuaris del sistema d'IA -amb "usuaris" no es refereix a les persones que veuen els seus drets afectats per un sistema d'IA, sinó a les que utilitzen el sistema d'IA¹¹⁸⁹, dins i fora la UE, sempre que afecti

¹¹⁸⁴ EBERS, M., "Standardizing AI. The Case of the European Commission's Proposal for an Artificial Intelligence Act", a DIMATTEO, L. A., CANNARSA, M. i PONCIBÒ, C. (eds.), *The Cambridge Handbook of Artificial Intelligence: Global Perspectives on Law and Ethics*, Cambridge, Cambridge University Press, 2022, pàg. 333.

¹¹⁸⁵ AÑÓN ROIG, M.J., "Desigualdades algorítmicas...", *op. cit.*, 2022, pàg. 24.

¹¹⁸⁶ EBERS, M., "Standardizing AI. The Case...", *op. cit.*, 2022, pàg. 325.

¹¹⁸⁷ Article 3.1 de la PRIA i Annex 1 de la PRIA. El Parlament Europeu esmena aquest article i proposa la definició de sistema d'IA com un sistema basat en una màquina que es designa per operar amb diversos nivells d'autonomia que pot, per objectius tan explícits com implícits, generar resultats com prediccions, recomanacions o decisions que influeixen entorns físics o virtuals.

¹¹⁸⁸ L'article 3.2 de la PRIA precisa que serà proveïdor tota persona física o jurídica, autoritat pública, agència o organisme d'una altra índole que desenvolupi un sistema d'IA o per el qual es va desenvolupar un sistema d'IA en vista d'introduir-lo en el mercat o posar-lo en servei amb el seu propi nom o marca comercial, ja sigui de manera remunerada o gratuïta. El considerant 53 de la PRIA estipula que convé que una persona física o jurídica concreta, definida com el proveïdor, assumeixi la responsabilitat associada a la introducció en el mercat o posada en servei d'un sistema d'IA de risc alt, amb independència de si aquesta persona física o jurídica és o no qui va dissenyar o desenvolupar el sistema.

¹¹⁸⁹ Tant l'article 3.4 com el considerant 59 de la PRIA estipulen que convé preveure que l'usuari d'un sistema d'IA ha de ser la persona física o jurídica, autoritat pública, agència o organisme d'un altre tipus sota l'autoritat de la qual s'utilitzi el sistema d'IA, llevat quan el seu ús s'emmarqui en una activitat personal de caràcter no professional.

persones situades en la UE¹¹⁹⁰. S'estableixen diferents tipus d'obligacions a qui proveeix els sistemes d'IA¹¹⁹¹ segons el risc (prohibits, risc alt, risc baix o mínim) que puguin causar a la sanitat, seguretat i drets fonamentals, de manera que, encara que es doni una definició àmplia d'IA, les obligacions dels proveïdors o usuaris es refereixen a subconjunts de sistemes que responen a definicions significativament més estretes¹¹⁹². Per exemple, pels sistemes d'IA de risc alt s'haurà d'establir un sistema de gestió de qualitat que garanteixi el compliment del Reglament¹¹⁹³, elaborar la documentació tècnica¹¹⁹⁴, assegurar que els seus sistemes són sotmesos al procés de conformitat oportú¹¹⁹⁵ i conservar els arxius de registre generats¹¹⁹⁶. També s'haurà de dur a terme una monitorització del sistema després de la seva implementació en el mercat¹¹⁹⁷, i si un sistema d'IA pot afectar negativament a la salut, seguretat o drets fonamentals s'haurà d'informar les autoritats nacionals competents¹¹⁹⁸.

Els sistemes de risc alt són els que protagonitzen la regulació, atès que els proveïdors que els comercialitzin hauran d'acreditar que compleixen, a través d'una avaluació de conformitat¹¹⁹⁹, uns requisits essencials -que es traduiran en estàndards harmonitzats creats per Organitzacions europees de normalització-, per obtenir el marcatge "CE" de conformitat i poder circular lliurement pel mercat intern de la UE. Es combinarà, pels sistemes de risc alt, aquesta avaluació de riscos *ex ante* amb una supervisió *ex post* per part de les autoritats de vigilància del mercat.

En aquest capítol, per estudiar la regulació de la IA des de la perspectiva de la no discriminació, en primer lloc, abordem quina rellevància se li ha donat al principi de no discriminació en la PRIA (1), en segon lloc estudiem la protecció contra la discriminació

¹¹⁹⁰ Article 2.1 de la PRIA. El considerant 10 estipula que les normes del Reglament s'apliquen a proveïdors de sistemes d'IA, amb independència de si estan establerts a la UE o en un país tercer, i als usuaris de sistemes d'IA establerts a la UE.

¹¹⁹¹ Article 16 de la PRIA. El Parlament Europeu en les seves esmenes al text de la Comissió proposa que les obligacions incloguin proporcionar especificacions sobre les dades d'entrada o qualsevol altra informació rellevant sobre les dades incloent les seves limitacions i assumpcions (...).

¹¹⁹² VEALE, M. i ZUIDERVEEN, F., "Demystifying the Draft EU Artificial Intelligence Act. Analysing the good, the bad, and the unclear elements of the proposed approach", *Computer Law Review International*, vol. 22, núm. 4, 2021, pàg. 109.

¹¹⁹³ Article 17 de la PRIA.

¹¹⁹⁴ Article 18 de la PRIA.

¹¹⁹⁵ Article 19 de la PRIA.

¹¹⁹⁶ Article 20 de la PRIA.

¹¹⁹⁷ Article 17.1.h i 61.1 de la PRIA.

¹¹⁹⁸ Article 22 i article 65 de la PRIA.

¹¹⁹⁹ Segons l'article 3.20 de la PRIA una avaluació de conformitat és el procés per el qual es verifica si es compleixen els requisits establerts en el Reglament en relació amb els sistemes d'IA de risc alt.

en cada nivell de risc que estableix la norma (2), en tercer lloc, quins elements podem extreure de la PRIA per protegir millor la discriminació algorítmica (3) i finalment, en quart lloc, com traslladar les exigències relatives al principi de no-discriminació als estàndards harmonitzats de la UE per obtenir el marcatge “CE”, i les exigències de control *ex post* dels sistemes d’IA de risc alt ja comercialitzats (4). Finalment, acabem el capítol amb un apartat de conclusions parcials (5).

1. El principi de no-discriminació en la PRIA

En la PRIA es motiva la perspectiva reguladora en gran part en base a la discriminació, és a dir, s’estipula que atès que els algorismes poden tenir efectes discriminatoris, han d’operar només sota certes condicions o directament ser prohibits (considerants 15, 17, 28, 33, 35-39, 44, 47)¹²⁰⁰. Cal puntualitzar que malgrat els considerants facin referència a la discriminació, no se’n fa de manera específica en l’articulat de la norma¹²⁰¹, si bé podem entendre que la no discriminació s’integra en un dret fonamental (concepte que sí que s’utilitza en nombroses ocasions en l’articulat). L’RGPD, en canvi, requereix que el tractament de dades personals es dugui a terme de tal manera que no produeixi efectes discriminatoris (considerant 71, 75, 85), sense tenir en compte l’existència d’efectes discriminatoris de manera tan directa. Tot i que les dues normes fan referència a la discriminació en més o menys mesura, cap de les dues defineix “discriminació” o “biaix” a la llum de particularitats específiques de la discriminació algorítmica, ni estipula principis específics o mesures per mitigar-la¹²⁰². A més, com bé apunta Berendt, tant l’RGPD como la PRIA operen en un context sociojurídic en el qual la discriminació de manera predominant és vista com a derivada d’una decisió concreta fent que l’enfocament tradicional que parteix de persones que prenen decisions que discriminen hagi portat a l’enfocament corresponent de màquines que prenen decisions que discriminen¹²⁰³. Alternativament a l’enfocament tradicional, i com ja hem defensat repetidament en aquest

¹²⁰⁰ BERENDT, B., “The AI Act Proposal...”, *op. cit.*, 2022, pàg. 34.

¹²⁰¹ El Parlament Europeu en el seu document d’esmenes a la proposta de Reglament de la Comissió sí que introdueix l’article 4 a relatiu a els principis generals aplicables a tots els sistemes d’IA, entre els quals es troba el principi de diversitat, no discriminació i equitat.

¹²⁰² *Ídem*, pàg. 35.

¹²⁰³ *Ídem*, pàg. 39-42.

treball, es pot considerar que tant les persones com les màquines responen a sistemes d'opressió que operen en la societat, donant rellevància a la discriminació de caire més sistèmic¹²⁰⁴.

Si bé es cert que, a diferència de l'RGPD la no discriminació cobra més protagonisme textualment, ens preguntem si això és únicament el que fa, motivar la regulació, reconèixer que els sistemes d'IA perpetuen patrons històrics de discriminació, però infravalorant les causes i l'abast de la discriminació algorítmica i donant poques solucions eficaces al respecte. Veale i Zuiderveen destaquen que la PRIA presenta problemes (per exemple de responsabilitat i compliment) que sorgeixen d'un enteniment massa estret dels límits dels sistemes que incorporen IA, les seves estructures i els rols dels actors involucrats¹²⁰⁵ (sobre aquest punt, vegeu capítol I).

Per entendre la posició del principi de no-discriminació en la regulació, cal tenir present que l'objectiu principal de la normativa proposada és assegurar el bon funcionament del mercat intern de la UE (1.1), fet que provoca que s'hagi tingut poca consideració de les persones afectades pels sistemes d'IA, no només en allò referent a la seva participació en la construcció de la regulació, sinó també en la introducció d'eines perquè de manera individual o col·lectiva es pugui fer front a incompliments del Reglament (1.2). A més, volem analitzar l'habilitació, en l'article 10.5 de la PRIA, a tractar categories de dades protegides per mitigar biaixos (1.3) i les implicacions d'una regulació basada en el risc pel principi de no-discriminació (1.4).

1.1. Un Reglament per assegurar el bon funcionament del mercat intern de la UE

L'objectiu general del Reglament és garantir un bon funcionament del mercat únic mitjançant les condicions necessàries pel desenvolupament i la utilització d'una

¹²⁰⁴ Vegeu, per exemple: BARRÈRE, M.A. i MORONDO, D., "Subordinación y discriminación interseccional...", *op. cit.*, 2011, pàg. 39.

¹²⁰⁵ BERENDT, B., "The AI Act Proposal...", *op. cit.*, 2022, pàg. 46. S'ha de destacar, així i tot, que el Reglament no fa distincions en els règims de protecció que presenta en base a si es tracta de processos completament automatitzats o en els quals intervenen persones, a diferència de l'article 22 de l'RGPD.

intel·ligència artificial fiable a la UE¹²⁰⁶. A la vegada persegueix quatre objectius específics, dos dels quals són garantir que els sistemes d'IA introduïts i emprats en el mercat de la UE siguin segurs i respectin la legislació vigent en matèria de drets fonamentals, i millorar la governança efectiva de la legislació vigent en matèria de drets fonamentals i els requisits de seguretat aplicables als sistemes d'IA¹²⁰⁷.

La base jurídica per proposar el Reglament és l'article 114 del TFUE, que tracta de l'adopció de mesures per garantir l'establiment i el funcionament del mercat interior¹²⁰⁸. La definició de les obligacions que s'aplicaran a proveïdors i usuaris que desenvolupin els sistemes i els comercialitzin al mercat de la UE, pretén generar seguretat jurídica i garantir que no sorgeixi cap obstacle per al subministrament transfronterer de serveis i productes associats a la IA¹²⁰⁹.

El considerant 1 de la PRIA estableix que es tracta de normes harmonitzades per l'ús de sistemes d'IA amb l'objectiu de prevenir que els estats membres imposin restriccions a l'ús d'intel·ligència artificial, si no és que el Reglament explícitament les autoritza. Per exemple, estableix que les obligacions a usuaris de sistemes d'IA de risc alt poden incrementar-se en la legislació de la UE o dels estats membres, però no hi ha cap provisió similar que s'apliqui a la totalitat del Reglament, de manera que existeix incertesa legal sobre si els requisits de transparència nacionals que s'apliquin més enllà de sistemes de risc alt, com per exemple les exigències de transparència algorítmica del sector públic a França¹²¹⁰, s'hauran de deixar d'aplicar¹²¹¹. Si aquest fos el cas, en comptes d'estar davant d'una regulació de drets mínims, estaríem davant d'una regulació de màxims.

Atès que l'objectiu principal del Reglament és assegurar el bon funcionament del mercat intern i s'ha adoptat una perspectiva basada en la seguretat dels productes, s'opta per una regulació de la intel·ligència artificial general en comptes de sectorial. S'ha argumentat contràriament que, com els riscos varien segons els sectors donat que diferents valors i

¹²⁰⁶ COMISIÓN EUROPEA, *Propuesta de Reglamento del Parlamento europeo y del Consejo, por el que se establecen normas armonizadas en materia de inteligencia artificial...*, *op. cit.*, 24 de abril de 2021, pàg. 10.

¹²⁰⁷ *Idem*, pàg. 3.

¹²⁰⁸ *Idem*, pàg. 6.

¹²⁰⁹ *Idem*, pàg. 12.

¹²¹⁰ Article 1 du Décret n° 2017-330 du 14 mars 2017 relatif aux droits des personnes faisant l'objet de décisions individuelles prises sur le fondement d'un traitement algorithmique. JORF n°0064 du 16 mars 2017. France.

¹²¹¹ VEALE, M. i ZUIDERVEEN, F., "Demystifying the Draft EU...", *op. cit.*, 2021, pàg. 109 i 110.

normes es posen en risc¹²¹² i l'ús dels sistemes d'intel·ligència artificial és massa variat per comptar només amb una estructura normativa¹²¹³, s'hauria d'apostar per una regulació sectorial (establint, per exemple, una protecció reforçada en sectors com el sanitari o l'educatiu). Una regulació sectorial, potser, també plantejaria normes més específiques que resoldrien alguns aspectes de la PRIA que indueixen a pensar que el principi d'equilibri que sembla inspirar tota la regulació no s'assoleix, ja que la perspectiva adoptada pel text és fonamentalment de gestió de riscos i que les garanties dels drets són insuficients¹²¹⁴.

Amb el Reglament es pretén compatibilitzar dos conceptes molt diferents. Garantir la lliure circulació transfronterera de béns i serveis basats en la IA, impedit als estats membres la imposició de restriccions al desenvolupament, comercialització i utilització de sistemes d'IA, a la vegada que s'assegura un nivell elevat de protecció de la salut, la seguretat i els drets humans¹²¹⁵. Nogensmenys, sembla difícil una promoció generalitzada de l'ús de sistemes d'IA que sigui compatible amb una perspectiva basada en els drets fonamentals. Com planteja Adams-Prassl, en referència a que s'hagi escollit una regulació de seguretat del producte, l'elecció d'una estructura normativa dissenyada per béns físics que poden causar danys tangibles és altament inapropiada per tecnologies digitals que poden suposar riscos significatius per a la protecció de valors intangibles com els drets fonamentals¹²¹⁶. Com apunta Roig Batalla, i per fer una comparació amb com ha operat aquest tipus de regulació en l'àmbit de la privadesa, s'han hagut de completar els requeriments de reducció de risc d'un producte amb exigències de sostenibilitat per incorporar valors socials i col·lectius que no es prenen en consideració en la definició de material segur¹²¹⁷.

Malgrat que en la PRIA s'utilitza un llenguatge orientat als drets fonamentals, la norma

¹²¹² ZUIDERVEEN, F., "Strengthening legal protection against discrimination..." *op. cit.*, 2020, pàg. 1586. L'autor argumenta que tal com no es va fer una normativa general quan va ocórrer la revolució industrial, tampoc s'hauria d'abordar la presa de decisions mitjançant sistemes algorítmics de manera general.

¹²¹³ ZUIDERVEEN, F., *Discrimination, artificial intelligence...*, *op. cit.*, 2018, pàg. 73.

¹²¹⁴ AÑÓN ROIG, M.J., "Desigualdades algorítmicas..." *op. cit.*, 2022, pàg. 28.

¹²¹⁵ Considerant 1 de la PRIA.

¹²¹⁶ ADAMS-PRASSL, J., "Regulating algorithms at work: Lessons for a 'European approach to artificial intelligence'", *European labour law journal*, vol.13, núm. 1, 2022, pàg. 49. L'autor planteja que el Reglament no sembla capaç d'abordar els riscos identificats en els seus propis considerants, no adreça les llacunes identificades entre diferents règims com la Protecció de Dades i l'antidiscriminació, i no clarifica terminologia important.

¹²¹⁷ ROIG, A., "Límits del principi de privadesa pel disseny", *Revista Catalana de Dret Públic*, núm. 64, 2022, pàg. 180.

prioritza facilitar i promoure el consum de sistemes d'IA (en base a la presumpció que es beneficiarà a totes les persones en tots els casos), de manera que pot ser que la protecció de drets fonamentals, com per exemple el dret a la no discriminació, es quedi en un àmbit més formal que de contingut i que serveixi únicament per motivar l'existència del Reglament. En aquest sentit, Edwards es mostra sorpresa que hi manqui una obligació *ex ante* d'una avaluació d'impacte dels drets fonamentals per tots els sistemes d'IA, en tant que normativa que pretén mitigar els riscos envers els drets fonamentals¹²¹⁸.

A continuació analitzem com, en seguir l'estructura d'una regulació de seguretat del producte, no se'ls hagi atorgat la consideració necessària a les persones afectades pels sistemes d'IA.

1.2. Poca consideració cap a les persones afectades pels sistemes d'IA

La proposta de la Comissió no confereix drets a les persones que es veuen afectades per sistemes d'IA, sense conferir mecanismes per desafiar o queixar-se dels sistemes d'IA quan aquests violen els seus drets¹²¹⁹. El Parlament Europeu, per la seva part, proposa que en el Reglament hi hagi una disposició que possibiliti el dret a les persones a presentar una queixa davant l'autoritat nacional de supervisió (art. 68.a i considerant 84.a), a un recurs judicial efectiu contra una autoritat de control (art. 68.b i considerant 84.a) i a rebre una explicació sobre una decisió individual basada en el resultat d'un sistema d'IA de risc alt quan produeixi efectes legals o d'afectació significativament similar (art. 68.c i considerant 84.b)¹²²⁰. És cert, igualment, que malgrat que no s'expliciti en la PRIA, els estats membres de la UE, han d'assegurar que les persones consumidores i altres parts interessades tenen l'oportunitat de generar queixes i que tinguin un seguiment adequat davant les autoritats de vigilància del mercat¹²²¹, malgrat que la legislació de la UE

¹²¹⁸ EDWARDS, L., *Regulating AI in Europe: four problems and four solutions*, Ada Lovelace Institute, London, March 2022, pàg. 13.

¹²¹⁹ EDWARDS, L., *Regulating AI in Europe: four problems...*, *op.cit.*, 2022, pàg. 10; VEALE, M. i ZUIDERVEEN, F., "Demystifying the Draft EU...", *op. cit.*, 2021, pàg. 111.

¹²²⁰ EUROPEAN PARLIAMENT, *Amendments adopted by the European Parliament on 14 June 2023...*, *op. cit.*, 14 June 2023.

¹²²¹ Comunicació de la Comissió, «Guía azul» sobre la aplicació de la normativa europea relativa a los productos, de 2022. DOUE C 247/1, 29.6.2022, pàg. 108.

simplement requereix que les autoritats de vigilància del mercat gestionin les queixes de manera competent i les considerin com qualsevol altra informació¹²²².

A més, s'afirma en l'exposició de motius que en cas que se segueixin produint violacions dels drets fonamentals malgrat el conjunt de proteccions *ex ante*, la transparència i traçabilitat garantides del sistema d'IA, unides a controls *ex post* sòlids, permetran oferir a les persones afectades una compensació efectiva¹²²³. Així i tot, no conté cap provisió per obtenir reparació individual¹²²⁴ o reparació col·lectiva (com la que confereix l'article 80 de l'RGPD). El Parlament Europeu, en aquest sentit, proposa afegir al text de la Comissió l'article 68.d que integra dins l'àmbit d'aplicació de la Directiva (UE) 2020/1828 del Parlamento Europeo y del Consejo, de 25 de noviembre de 2020, relativa a las acciones de representación para la protección de los intereses colectivos de los consumidores¹²²⁵ (d'ara en endavant, Directiva 2020/1828) el Reglament de la IA per tal que sigui una matèria que pugui ser objecte d'una acció de representació¹²²⁶. La proposta de la Comissió tampoc presenta cap mecanisme a través del qual les persones o la societat civil puguin participar en els processos d'investigació de sistemes d'IA de risc alt, de manera que no adreça el conjunt de perjudicis que poden sorgir de l'opacitat, complexitat, escala i asimetria de poder en les quals es desenvolupen els sistemes d'IA¹²²⁷. L'absència de consideració a les persones i comunitats afectades ha estat criticada pel Comitè Europeu de Protecció de Dades i el Supervisor Europeu de Protecció de dades, que precisen que tant si es tracta d'usuaris finals, com de persones interessades o persones

¹²²² Article 11.3.e i 11.7.a del Reglamento (UE) 2019/1020, del Parlamento Europeo y del Consejo, de 20 de junio de 2019, relativo a la vigilancia del mercado y la conformidad de los productos y por el que se modifican la Directiva 2004/42/CE y los Reglamentos (CE) núm. 765/2008 y (UE) núm. 305/2011. DOUE, L 169/1, 25.6.2019.

¹²²³ COMISIÓN EUROPEA, *Propuesta de Reglamento del Parlamento europeo y del Consejo, por el que se establecen normas armonizadas en materia de inteligencia artificial...*, *op. cit.*, 24 de abril de 2021, pàg. 13.

¹²²⁴ SMUHA, N. A., "Beyond the individual: governing AI's societal harm", *Internet Policy Review*, vol. 10, núm. 3, 2021, pàg. 22. L'autora especifica que en el Reglament no es dona cap dret d'accés a la justícia directe, d'accés a la informació o de participació.

¹²²⁵ Directiva (UE) 2020/1828 del Parlamento Europeo y del Consejo, de 25 de noviembre de 2020, relativa a las acciones de representación para la protección de los intereses colectivos de los consumidores, y por la que se deroga la Directiva 2009/22. DOUE, L 409/1, 4.12.2020.

¹²²⁶ EUROPEAN PARLIAMENT, *Amendments adopted by the European Parliament on 14 June 2023...*, *op. cit.*, 14 June 2023.

¹²²⁷ EDRI *et al.*, "An EU Artificial Intelligence Act for Fundamental Rights. A civil Society Statement", 30 November 2021, pàg. 4. Disponible a: <https://edri.org/wp-content/uploads/2021/12/Political-statement-on-AI-Act.pdf> Darrera visita el 2 de setembre de 2023.

afectades pel sistema d'IA, la seva absència en el text suposa un punt cec de la regulació¹²²⁸.

En l'article 13 de la PRIA, s'estableixen obligacions de transparència i comunicació a usuaris (que no són usuaris finals)¹²²⁹ quan es tracti de sistemes de risc alt. En aquesta informació ha de constar, entre altres ítems que no són tan rellevants quan ens referim a la discriminació algorítmica, el funcionament del sistema d'IA en referència a les persones o grups de persones en relació amb els quals es pretengui emprar el sistema¹²³⁰. El Parlament Europeu en les seves esmenes a la proposta de la Comissió afegeix, entre la informació que ha de constar el grau en el qual el sistema d'IA pot proveir una explicació per les decisions que pren¹²³¹. Tant el funcionament del sistema com el grau en el qual el sistema d'IA pot proveir una explicació per les decisions que pren interessaria en el cas d'iniciar un plet de discriminació, atès que hi hauria unes bases per evidenciar si el sistema, per al grup de població on es troba cada persona, és més o menys acurat i quina incidència pot haver tingut aquesta manca de precisió en què després s'hagi donat un resultat discriminatori. Però, aquesta informació no estarà disponible per a les persones usuàries finals, ja que aquest article està orientat a dissenyar i desenvolupar els sistemes d'IA d'una manera que garanteixi que funcionen amb un nivell de transparència suficient per tal que els usuaris interpretin i usin correctament la seva informació de sortida¹²³². No obstant això, a Espanya, Cotino Hueso ha plantejat que la documentació tècnica -que obliga a confeccionar la PRIA pels sistemes de risc alt- que estigui en possessió de les autoritats de supervisió, en molts casos -sobretot quan es tracti d'usos públics de la IA- es podrà considerar informació pública¹²³³ a la qual les persones podran accedir a través del dret d'accés a informació pública¹²³⁴.

¹²²⁸ EUROPEAN DATA PROTECTION BOARD i EUROPEAN DATA PROTECTION SUPERVISOR, *Joint Opinion 5/2021 on the Proposal for a Regulation of the European Parliament and of the Council Laying down Harmonised Rules on Artificial Intelligence (Artificial Intelligence Act)*, Brussels, 18 June 2021, par. 18.

¹²²⁹ En la proposta de la Comissió s'utilitza el terme "user", i en canvi, el Parlament Europeu en el seu document d'esmenes a la proposta de la Comissió empra el terme "deployer".

¹²³⁰ Article 13.3.b.iv de la PRIA.

¹²³¹ EUROPEAN PARLIAMENT, *Amendments adopted by the European Parliament on 14 June 2023...*, *op. cit.*, 14 June 2023.

¹²³² Article 13 de la PRIA.

¹²³³ L'article 13 de la Llei 19/2013 disposa que: *Se entiende por información pública los contenidos o documentos, cualquiera que sea su formato o soporte, que obren en poder de alguno de los sujetos incluidos en el ámbito de aplicación de este título y que hayan sido elaborados o adquiridos en el ejercicio de sus funciones.*

¹²³⁴ COTINO HUESO, L. "Qué concreta transparencia e información de algoritmos e inteligencia artificial es la debida", *Revista Española de la Transparencia (RET)*, núm. 16, 2023, pàg. 57.

Per últim, les úniques obligacions de les persones usuàries de sistemes d'IA envers les persones usuàries finals que trobem en la PRIA són relatives a la transparència¹²³⁵. Com desenvolupem en l'apartat 2.3 d'aquest capítol, aquestes obligacions es basen en informar les persones que estan interactuant amb un sistema d'intel·ligència artificial quan aquest té la capacitat, per exemple, de reconèixer les emocions o realitzar categorització biomètrica¹²³⁶.

Per a la majoria de sistemes d'IA de risc alt, però, aquesta transparència té un abast limitat (sobretot per la manca de coneixement especialitzat de les persones per entendre, establir una posició crítica i desafiar sistemes d'IA perjudicials i opacs). En la PRIA, es confia en la transparència com a contra-mesura, sense tenir en compte el coneixement científic existent de les maneres a través de les quals els sistemes d'IA, com a parts de sistemes sociotècnics més amplis, creen efectes discriminatoris¹²³⁷. Una altra mesura en la qual confia és en l'habilitació del tractament de categories protegides amb la finalitat de mitigar biaixos, que estudiem tot seguit.

1.3. Habilitació del tractament de categories especials de dades

L'article 10.5 de la PRIA, estableix que en la mesura en què sigui estrictament necessari per garantir la vigilància, detecció i correcció dels biaixos associats al sistema d'IA de risc alt, les persones proveïdores dels sistemes podran tractar les categories especials de dades personals que es mencionen en l'article 9.1 de l'RGPD, l'article 10 de la Directiva 2016/680 i l'article 10 del Reglamento (UE) 2018/1725 del Parlamento Europeo y del Consejo, de 23 de octubre de 2018, relativo a la protección de las personas físicas en lo que respecta al tratamiento de datos personales por las instituciones, órganos y

¹²³⁵ Article 52 de la PRIA. L'article 13 del Reglament, en canvi, imposa obligacions de transparència a les persones proveïdores enfront les persones usuàries, però sense establir criteris de transparència enfront les persones afectades.

¹²³⁶ Article 52.2 de la PRIA. No s'aplicarà si estan autoritzats per la llei amb l'objectiu de prevenir infraccions penals.

¹²³⁷ BERENDT, B., "The AI Act Proposal...", op. cit., 2022, pàg. 32. L'autora també opina que la transparència, malgrat que estigui aparellada amb la qualitat de les dades, seguretat, supervisió humana i documentació, no és suficient per prevenir la discriminació.

organismos de la Unión, y a la libre circulación de esos datos¹²³⁸, oferint les salvaguardes adequades pels drets i les llibertats fonamentals de les persones físiques, que inclouen establir limitacions tècniques a la reutilització i la utilització de mesures de seguretat i protecció de la privacitat més recents, tals com la pseudonimització o el xifratge, quan l'anonimització pugui afectar significativament l'objectiu perseguit¹²³⁹.

Aquest article vincula la detecció de biaixos associats al sistema d'IA amb el tractament de dades protegides especialment per la legislació. En la PRIA, sembla que s'ha fet cas de les reclamacions per part de l'acadèmia i de la societat civil que alertaven sobre la dificultat de detectar el potencial dels sistemes d'IA per causar discriminació indirecta quan el sistema algorítmic desconeix les categories de dades protegides de les persones i comunitats afectades¹²⁴⁰.

L'article 9.2 de l'RGPD també permet el tractament de categories protegides de dades per diverses raons, com per exemple, per un interès públic essencial, sobre la base del dret de la UE o dels estats membres¹²⁴¹, però no per raó de detecció de biaixos¹²⁴². Dins les circumstàncies que permeten el tractament de categories especialment protegides en l'RGPD n'hi ha algunes que afegeixen la necessitat d'establir mesures suplementàries. La necessitat per raó d'interès públic essencial és la que disposa mesures més exigents, ja que s'estableix que el tractament de dades personals especialment protegides, ha de ser proporcional a l'objectiu perseguit, respectar en allò essencial el dret a la protecció de dades i establir mesures adequades i específiques per protegir els interessos i drets

¹²³⁸ Reglamento (UE) 2018/1725 del Parlamento Europeo y del Consejo, de 23 de octubre de 2018, relativo a la protección de las personas físicas en lo que respecta al tratamiento de datos personales por las instituciones, órganos y organismos de la Unión, y a la libre circulación de esos datos, y por el que se derogan el Reglamento (CE) nº 45/2001 y la Decisión nº 1247/2002/CE. DOUE, L 295/39, 21/11/2018.

¹²³⁹ Article 10.5 de la PRIA.

¹²⁴⁰ Vegeu, per exemple: HOFFMANN, H., VOGT, V., HAUER, M. P. i ZWEIG, K., "Fairness by awareness? On the inclusion of protected...", *op. cit.*, 2022, pàg. 11; VEALE, M. i BINNS, R., "Fairer machine learning in the realworld: Mitigating discrimination without collecting sensitive data", *Big Data & Society*, vol. 4, núm. 2, 2017, pàg. 4. En l'àmbit espanyol, vegeu: SORIANO ARNANZ, A., "Decisiones automatizadas y discriminación...", *op. cit.*, 2021, pàg. 38; SORIANO ARNANZ, A., "La propuesta de Reglamento..." *op. cit.*, 2021, pàg. 16; SORIANO ARNANZ, A., "Creating non-discriminatory Artificial Intelligence systems: balancing the tensions between code granularity and the general nature of legal rules", *Revista d'Internet, Dret i Política*, núm. 38, 2023, pàg. 9; COTINO HUESO, L., "Discriminación, sesgos e igualdad de la inteligencia artificial...", *op. cit.*, 2023, pàg. 300.

¹²⁴¹ Article 9.2 de l'RGPD.

¹²⁴² VAN BEKKUM, M. i ZUIDERVEEN, F., "Using sensitive data to prevent discrimination by artificial intelligence: Does the GDPR need a new exception?" *Computer Law and Security Review*, vol. 48, 2023, pàg. 1-12. Els autors d'aquest article analitzen la necessitat o no de crear una excepció a la prohibició de tractament de dades de l'article 9 de l'RGPD per mitigar els biaixos. Analitzen entre altres coses, les salvaguardes plantejades en l'article 10.5 de la PRIA.

fonamentals de la persona interessada¹²⁴³. L'article 10.5 de la PRIA, en canvi, no menciona explícitament la proporcionalitat i la protecció de la persona sembla ser més dèbil (menciona “*mesures apropiades*” en comptes de “*mesures adequades i específiques per protegir*”)¹²⁴⁴. Tanmateix, en les esmenes del Parlament Europeu sobre el text presentat per la Comissió sí que concreta certes mesures a aplicar com l'accés a les dades només per part de persones autoritzades o la creació de documentació que justifiqui la necessitat del tractament de categories especials de dades¹²⁴⁵.

L'exempció de l'article 10.5 de la PRIA està basada en necessitat estricta i està subjecte a certes salvaguardes, només es pot usar per sistemes de risc alt i només pels proveïdors dels sistemes, de manera que els proveïdors de sistemes que no siguin de risc alt no poden emprar-la (no obstant això, en l'article 69 de la PRIA s'encoratja la creació de codis de conducta per aplicar els requisits essencials a tots els sistemes d'IA, independentment del grau de risc que se'ls atorgui)¹²⁴⁶.

El tractament de categories protegides de dades pot suposar una interferència en la privacitat de les persones, pot suposar l'abús en la utilització de les dades o la utilització per altres usos aliens a la detecció de la discriminació¹²⁴⁷. Per això, especifica el considerant 41 de la PRIA, que tot ús d'un sistema d'IA de risc alt ha de dur-se a terme en consonància amb els requisits oportuns derivats de la CDFUE i els actes aplicables del dret derivat de la UE i el dret intern, per tant, no s'ha d'entendre que constitueix el fonament jurídic per al tractament de dades personals, incloses categories especials de dades, quan sigui pertinent¹²⁴⁸.

En aquest sentit, apunten Balayn i Gürses, que paradoxalment es possibilita l'ús de

¹²⁴³ Article 9.2.g de l'RGPD.

¹²⁴⁴ BERENDT, B., “The AI Act Proposal...”, *op. cit.*, 2022, pàg. 50.

¹²⁴⁵ EUROPEAN PARLIAMENT, *Amendments adopted by the European Parliament on 14 June 2023...*, *op. cit.*, 14 June 2023.

¹²⁴⁶ VEALE, M. i ZUIDERVEEN, F., “Demystifying the Draft EU...”, *op. cit.*, 2021, pàg. 103. En el considerant 81 del Reglament es precisa que s'ha d'animar als proveïdors de sistemes d'IA que no són de risc alt a crear codis de conducta i aplicar, amb caràcter voluntari, requisits addicionals relatius, per exemple, a la sostenibilitat mediambiental, l'accessibilitat per les persones amb discapacitat, la participació de les parts interessades en el disseny i desenvolupament de sistemes d'IA, i diversitat en els equips de desenvolupament del sistema. Cal puntualitzar que el Parlament Europeu en les seves esmenes al text presentat per la Comissió planteja que els codis de conducta, entre altres, han de c) considerar la manera en la qual l'ús dels sistemes d'IA tindrà un impacte podrà incrementar la diversitat, l'equilibri de gènere i la igualtat i d) tenir en compte si els sistemes d'IA poden ser usats de manera que, directament o indirectament, poden residualment o significativament, reforçar biaixos o desigualtats existents.

¹²⁴⁷ VAN BEKKUM, M. i ZUIDERVEEN, F., “Using sensitive data...”, *op. cit.*, 2023, pàg. 12.

¹²⁴⁸ Considerant 41 de la PRIA.

categories protegides per la detecció de biaixos sense considerar els riscos que imposa en les poblacions que la normativa presumptivament vol protegir¹²⁴⁹. Van Bekkum i Zuiderveen afegeixen que permetre les organitzacions recollir categories especials de dades no garanteix que les organitzacions puguin des-esbiaixar els seus sistemes d'IA¹²⁵⁰. Tot seguit estudiem l'adequació de la perspectiva basada en el risc en relació amb el principi de no-discriminació.

1.4. Una perspectiva basada en el risc

S'ha argumentat que no tots els usos dels algoritmes en la presa de decisions han de desencadenar un règim de regulació complet i que la presa de decisions mitjançant algoritmes pot regular-se de manera diferent o a diferents escales, en diferents contextos polítics i àrees del dret¹²⁵¹. Seguint aquesta perspectiva reguladora, la PRIA planteja un enfocament basat en el risc que només prohibeix els sistemes d'IA o imposa càrregues normatives, quan és probable que un sistema d'IA suposi riscos alts pels drets fonamentals i la seguretat¹²⁵². Sota aquesta perspectiva, només quan el risc pels drets fonamentals per l'aplicació d'un sistema d'IA es considera alt, es requereix una resposta normativa¹²⁵³. Per determinar si un sistema és o no de risc alt és particularment pertinent, segons el considerant 28 de la PRIA, la magnitud de les conseqüències adverses d'un sistema d'IA pels drets fonamentals protegits a la CDFUE¹²⁵⁴. El dret a la no discriminació és un dret, com ja hem esmentat, inclòs en la CDFUE, de manera que el risc que un sistema d'IA causi discriminació algorítmica hauria de ser particularment pertinent en la determinació del grau de risc del sistema.

La perspectiva *ex ante* que designa els sistemes d'IA a diferents categories de risc no considera que el nivell de risc depèn també del context en el qual el sistema es

¹²⁴⁹BALAYN, A. i GÜRSES, S., *Beyond debiasing. Regulating AI and Its Inequalities*, European Digital Rights (EDRi), Brussels, 2021, pàg. 94.

¹²⁵⁰ VAN BEKKUM, M. i ZUIDERVEEN, F., "Using sensitive data...", *op. cit.*, 2023, pàg. 12.

¹²⁵¹ KAMINSKI, M.E., "Binary Governance...", *op. cit.*, 2019, pàg. 1551.

¹²⁵² COMISIÓN EUROPEA, *Propuesta de Reglamento del Parlamento europeo y del Consejo, por el que se establecen normas armonizadas en materia de inteligencia artificial...*, *op. cit.*, 24 de abril de 2021, pàg. 7.

¹²⁵³ GUIHOT, M. *et al.*, "Nudging Robots: Innovative...", *op. cit.*, 2017, pàg. 452.

¹²⁵⁴ ZUDDAS, P., "Decisión algorítmica y discriminación", *op. cit.*, 2022, pàg. 10.

desenvolupa i no pot ser completament determinat d'avançada¹²⁵⁵. En aquest sentit, no obstant que, per exemple, en el cas del risc que suposa la discriminació algorítmica s'estableixi un marc normatiu de tutela preventiva¹²⁵⁶, aquest no és suficient quan la discriminació algorítmica depèn estretament també del context en el qual s'aplica el sistema d'IA. També, la classificació segons el risc genera que els sistemes d'IA no inclosos ni en aquells prohibits ni en aquells de risc alt, quedin fora de les normes dirigides a reduir els riscos pels drets fonamentals, passant d'algunes normes a pràcticament cap norma¹²⁵⁷. A més, aquesta classificació, per ser el més socialment precisa possible hauria de ser fruit de processos democràtics i inclusivament per determinar les prohibicions i les aplicacions de risc alt, incloent la societat civil i les comunitats afectades. Com ja hem esmentat, no sembla que aquesta inclusió s'hagi dut a terme de manera significativa. Com estableix Guidot *et al.*¹²⁵⁸, els riscos només poden ser degudament avaluats amb tota la informació pertinent, i això comporta tant aquelles persones que es troben en la indústria com totes les parts rellevants interessades. Finalment, Edwards¹²⁵⁹ arriba a afirmar que la naturalesa basada en el risc de la PRIA és il·lusòria i arbitrària, atès que l'impacte en grups determinats de la societat i la societat com a conjunt, ha de ser considerat igual que els riscos per a les persones i els seus drets i que, com ja hem esmentat, el risc ha de ser considerat durant tot el cicle de vida del sistema d'IA, no només en la seva entrada en el mercat.

Recapitulant, l'enfocament que planteja la PRIA per regular els sistemes d'IA a través d'una perspectiva de seguretat en el producte col·lideix amb la seva pretesa vocació de protecció genèrica dels drets fonamentals (entre ells la no discriminació). No s'atorguen drets a les persones interessades (malgrat que això pot canviar en el text final arran de les esmenes proposades per part del Parlament Europeu¹²⁶⁰) i la classificació del risc de cada sistema s'estableix d'avançada, fet problemàtic atès que en la discriminació algorítmica el nivell de risc depèn també del context en el qual el sistema es desenvolupa. Tanmateix, l'habilitació del tractament de categories especials de dades per garantir la vigilància,

¹²⁵⁵ EDRI *et al.*, "An EU Artificial Intelligence Act for Fundamental Rights. A civil Society Statement", 30 November 2021, pàg. 1. Disponible a: <https://edri.org/wp-content/uploads/2021/12/Political-statement-on-AI-Act.pdf> Darrera visita el 2 de setembre de 2023.

¹²⁵⁶ SÁEZ LARA, C., "El algoritmo como protagonista de la relación laboral...", *op. cit.*, 2020, pàg. 58.

¹²⁵⁷ VEALE, M. i ZUIDERVEEN, F., "Demystifying the Draft EU...", *op. cit.*, 2021, pàg. 110.

¹²⁵⁸ GUIHOT, M. *et al.*, "Nudging Robots: Innovative...", *op. cit.*, 2017, pàg. 451.

¹²⁵⁹ EDWARDS, L., *Regulating AI in Europe: four problems...*, *op. cit.*, 2022, pàg. 11.

¹²⁶⁰ EUROPEAN PARLIAMENT, *Amendments adopted by the European Parliament on 14 June 2023...*, *op. cit.*, 14 June 2023.

detecció i correcció dels biaixos associats al sistema d'IA de risc alt sembla una bona proposta, sempre que s'estableixin salvaguardes reals i efectives i no s'acabi desprotegint els col·lectius més vulnerables. A continuació, analitzem, més específicament les classificacions basades en el risc.

2. Protecció de la discriminació algorítmica segons el grup de risc en el qual estigui classificat el sistema d'IA

L'enfocament basat en el risc divideix els usos de sistemes de la IA entre prohibits (2.1), amb un risc alt (2.2), o amb un risc baix o mínim¹²⁶¹(2.3). En aquest apartat pretenem analitzar la protecció envers la discriminació en cada nivell de risc.

2.1. Sistemes prohibits

La PRIA estableix en l'article 5.1, quatre categories prohibides de sistemes d'IA. Les tres primeres es prohibeixen completament, dues en base a la manipulació i una en base a la qualificació social, i en canvi, l'última en base a la identificació biomètrica en temps real i remot, permet alguna aplicació.

En primer lloc, es prohibeix la introducció en el mercat, posada en servei o ús dels sistemes d'IA que es serveixin de tècniques subliminals que transcendeixen la consciència d'una persona per alterar de manera substancial el seu comportament provocant o essent altament probable que provoquin perjudicis físics o psicològics a aquesta persona o a una altra¹²⁶². Amb aquesta prohibició s'exclou la comercialització de productes que produeixin qualificacions o recomanacions discriminatòries en mercats en línia¹²⁶³.

¹²⁶¹ COMISIÓN EUROPEA, *Propuesta de Reglamento del Parlamento europeo y del Consejo, por el que se establecen normas armonizadas en materia de inteligencia artificial...*, *op. cit.*, 24 de abril de 2021, pàg. 14.

¹²⁶² Article 5.1.a de la PRIA.

¹²⁶³ VEALE, M. i ZUIDERVEEN, F., "Demystifying the Draft EU...", *op. cit.*, 2021, pàg. 99.

En segon lloc, es prohibeix la introducció en el mercat, posada en servei o utilització d'un sistema d'IA que aprofiti alguna de les vulnerabilitats d'un grup específic de persones deguda la seva edat o discapacitat física o mental per alterar de manera substancial el comportament d'una persona que pertanyi a aquest grup, provocant o essent altament probable que provoquin perjudicis físics o psicològics a aquesta persona o a una altra¹²⁶⁴. S'ha remarcat la necessitat d'expandir aquest article per incloure un conjunt complet de vulnerabilitats, en comptes de limitar-lo a l'edat i la discapacitat física o mental. L'article hauria de prohibir l'ús de qualsevol sistema d'IA que exploti les vulnerabilitats d'una persona o grup basades en qualsevol característica protegida especialment per la legislació¹²⁶⁵.

En l'exposició de motius es planteja que la legislació vigent en matèria de protecció de dades, protecció de les persones consumidores i de serveis digitals, que segons la PRIA, garanteixen que les persones físiques siguin degudament informades i puguin decidir lliurement no ser sotmeses a l'elaboració de perfils o altres pràctiques que puguin afectar la seva conducta, podrien cobrir altres pràctiques de manipulació o d'explotació contra adults que els sistemes d'IA poden facilitar¹²⁶⁶. Aquestes dues provisions, que es refereixen a la manipulació, tenen com a objectiu prohibir aquells sistemes d'IA comercialitzats i emprats per distorsionar el comportament humà. No sembla probable que els proveïdors admetin aquesta intenció, i per tant, pot ser que aquestes dues provisions tinguin un impacte pràctic reduït. La reconfiguració de la IA és una tendència significativa. Actualment ja es comú disfressar els productes del mercat, per exemple, el producte “stalkerware”, que és un software espia que algú instal·la en un dispositiu amb l'objectiu de vigilar i controlar de manera continua i que es ven com a rastrejador d'infants¹²⁶⁷.

En tercer lloc, es prohibeix la introducció en el mercat, posada en servei o utilització d'un sistema d'IA per part d'autoritats públiques o en la seva representació, amb la finalitat d'avaluar o classificar la fiabilitat de persones físiques durant un període determinat de

¹²⁶⁴ Article 5.1.b de la PRIA.

¹²⁶⁵ EDRI *et al.*, “An EU Artificial Intelligence Act for Fundamental Rights. A civil Society Statement”, 30 November 2021, pàg. 2. Disponible a: <https://edri.org/wp-content/uploads/2021/12/Political-statement-on-AI-Act.pdf> Darrera visita el 2 de setembre de 2023.

¹²⁶⁶ COMISIÓ EUROPEA, *Propuesta de Reglamento del Parlamento europeo y del Consejo, por el que se establecen normas armonizadas en materia de inteligencia artificial...*, *op. cit.*, 24 de abril de 2021, pàg. 14 i 15.

¹²⁶⁷ VEALE, M. i ZUIDERVEEN, F., “Demystifying the Draft EU...”, *op. cit.*, 2021, pàg. 100.

temps, atenent a la seva conducta social o a característiques personals o de la seva personalitat conegudes o predites, de forma que la qualificació social resultant provoqui, o bé un tracte perjudicial o desfavorable cap a determinades persones físiques o col·lectius sencers en contextos socials que no guarden relació amb els contextos on es van generar o recopilar les dades originalment, o bé un tracte perjudicial o desfavorable cap a determinades persones físiques o col·lectius sencers que és injustificat o desproporcionat respecte al seu comportament social o la gravetat d'aquest¹²⁶⁸. La prohibició d'aquests sistemes d'IA ve determinada per la seva capacitat per tenir efectes discriminatoris i abocar a l'exclusió de determinats grups, menystenint el dret a la dignitat i la no discriminació i els valors d'igualtat i justícia¹²⁶⁹. Es destaca positivament la regulació explícita de sistemes d'IA que porten a un tracte perjudicial o desfavorable de certes persones o grups, però cal destacar que ho fa en un context relativament estret de sistemes d'IA prohibits per l'article 5 de la PRIA¹²⁷⁰, quan molts sistemes poden generar aquest tipus de resultat sense ser la seva intenció principal.

En quart lloc, es prohibeix l'ús de sistemes d'identificació biomètrica “en temps real” en espais d'accés públic amb finalitats d'aplicació de la llei, llevat i en la mesura en què el seu ús sigui estrictament necessari per a la recerca selectiva de víctimes d'un delictes, la prevenció d'una amenaça imminent o atac terrorista, o la localització, identificació o detenció d'una persona que ha comès o presumptivament ha comès un delictes de l'article 2.2 de la Decisión Marco 2002/584/JAI del Consejo de 13 de junio de 2002 relativa a la orden de detención europea y a los procedimientos de entrega entre Estados miembros¹²⁷¹.

¹²⁶⁸ Article 5.1.c de la PRIA.

¹²⁶⁹ Considerant 17 de la PRIA.

¹²⁷⁰ BERENDT, B., “The AI Act Proposal...”, op. cit., 2022, pàg. 37.

¹²⁷¹ Article 5.1.d de la PRIA. L'article 2.2 de la Decisión Marco 2002/584/JAI del Consejo de 13 de junio de 2002 relativa a la orden de detención europea y a los procedimientos de entrega entre Estados miembros, inclou els següents delictes: pertenencia a organización delictiva, terrorismo, trata de seres humanos, explotación sexual de los niños y pornografía infantil, tráfico ilícito de estupefacientes y sustancias psicotrópicas, tráfico ilícito de armas, municiones y explosivos, corrupción, fraude, incluido el que afecte a los intereses financieros de las Comunidades Europeas con arreglo al Convenio de 26 de julio de 1995 relativo a la protección de los intereses financieros de las Comunidades Europeas, blanqueo del producto del delito, falsificación de moneda, incluida la falsificación del euro, delitos de alta tecnología, en particular delito informático, delitos contra el medio ambiente, incluido el tráfico ilícito de especies animales protegidas y de especies y variedades vegetales protegidas, ayuda a la entrada y residencia en situación ilegal, homicidio voluntario, agresión con lesiones graves, tráfico ilícito de órganos y tejidos humanos, secuestro, detención ilegal y toma de rehenes, racismo y xenofobia, robos organizados o a mano armada, tráfico ilícito de bienes culturales, incluidas las antigüedades y las obras de arte, estafa, chantaje y extorsión de fondos, violación de derechos de propiedad industrial y falsificación de mercancías, falsificación de documentos administrativos y tráfico de documentos falsos, falsificación de medios de pago, tráfico ilícito de sustancias hormonales y otros factores de crecimiento, tráfico ilícito de materiales radiactivos o

El Parlament Europeu proposa que el text d'esmenes a la proposta de la Comissió que es prohibeixi qualsevol ús de sistemes d'identificació biomètrica “en temps real” en espais d'accés públic i a més proposa prohibir també l'ús de sistemes de categorització biomètrica que categoritzen persones naturals segons atributs sensibles o protegits o la inferència d'aquests¹²⁷².

El considerant 33 de la norma ja especifica que les imprecisions tècniques dels sistemes d'IA destinats a la identificació biomètrica remota de les persones físiques pot donar lloc a resultats esbiaixats i tenir conseqüències discriminatòries. Això és especialment important en allò que respecta a l'edat, l'ètnia, el sexe o la discapacitat¹²⁷³. Només els sistemes en temps real que capturin, comparin i identifiquin de manera instantània, gairebé instantània o sense dilació significativa estan prohibits. Això exclou sistemes que, per exemple, analitzen gravacions de vídeo biomètricament després d'un esdeveniment, per exemple, per identificar persones després d'una manifestació, i sistemes que categoritzen persones biomètricament¹²⁷⁴.

Es defensa que més que prohibir, en la PRIA es legitima la vigilància de la població a gran escala. El Comitè Europeu de Protecció de Dades i els Supervisor Europeu de Protecció de Dades demanen una prohibició general de l'ús d'IA de reconeixement automàtic de trets humans en espais accessibles al públic en qualsevol context¹²⁷⁵.

2.2. Sistemes de risc alt

En la PRIA, a part dels sistemes prohibits, també s'estableixen uns requisits mínims necessaris per esmenar els riscos i problemes vinculats a la IA, establint criteris per definir

sustancias nucleares, tráfico de vehículos robados, violación, incendio voluntario, delitos incluidos en la jurisdicción de la Corte Penal Internacional, secuestro de aeronaves y buques, sabotaje.

¹²⁷² EUROPEAN PARLIAMENT, *Amendments adopted by the European Parliament on 14 June 2023...*, *op. cit.*, 14 June 2023.

¹²⁷³ Considerant 33 de la PRIA.

¹²⁷⁴ VEALE, M. i ZUIDERVEEN, F., “Demystifying the Draft EU...”, *op. cit.*, 2021, pàg. 101. El Parlament Europeu en el text on proposa esmenes a la proposta de la Comissió sí que planteja que també es prohibeixin els sistemes d'anàlisi biomètrica d'imatges gravades en espais públic posteriorment a que hagin estat gravades.

¹²⁷⁵ EUROPEAN DATA PROTECTION BOARD i EUROPEAN DATA PROTECTION SUPERVISOR, *Joint Opinion 5/2021 on the Proposal...*, *op. cit.*, 18 June 2021, par. 31.

aquells sistemes d'IA que plantegen un risc alt per a la salut i la seguretat o els drets fonamentals de les persones. Aquests sistemes d'IA hauran de complir una sèrie de requisits horitzontals obligatoris que, en teoria, garanteixin la seva fiabilitat, i ser sotmesos a procediments d'avaluació de la conformitat abans de poder introduir-se en el mercat de la UE¹²⁷⁶.

La pretensió és establir normes comunes per a tots els sistemes d'IA de risc alt amb l'objectiu de garantir un nivell elevat i coherent de protecció dels interessos públics en allò que respecta a la salut, la seguretat i els drets fonamentals. Les normes han de ser coherents amb la CDFUE, no han de ser discriminatòries i han d'estar en consonància amb els compromisos de la UE en matèria de comerç internacional¹²⁷⁷.

En el capítol I del Títol III de la PRIA (títol de la norma que es refereix íntegrament als sistemes de risc alt), específicament en l'article 6, s'estableixen les normes de classificació i es defineixen les dues categories principals de sistemes d'IA de risc alt, a saber, els sistemes d'IA dissenyats per utilitzar-se com a components de seguretat de productes, subjectes a una avaluació de conformitat *ex ante* realitzada per tercers i enumerats en l'annex II de la norma (per exemple, màquines, jocs o ascensors), i altres sistemes d'IA independents, amb implicacions relacionades principalment amb els drets fonamentals (els quals s'indiquen en l'annex III)¹²⁷⁸. Interessa apuntar que la mateixa norma estableix que la qualificació de risc alt s'ha de limitar a aquells sistemes d'IA que tinguin conseqüències perjudicials importants per a la salut, la seguretat i els drets fonamentals de les persones de la UE¹²⁷⁹. Per tant, no pretén fer front a qualsevol conseqüència, sinó a conseqüències *importantes*, sense establir-se cap criteri sobre que es considera o no una conseqüència important.

En l'annex III, s'especifica que es consideren de risc alt els sistemes d'IA que s'empren en els àmbits següents: identificació biomètrica i categorització de persones¹²⁸⁰; gestió i

¹²⁷⁶ COMISIÓN EUROPEA, *Propuesta de Reglamento del Parlamento europeo y del Consejo, por el que se establecen normas armonizadas en materia de inteligencia artificial...*, op. cit., 24 de abril de 2021, pàg. 3 i 4.

¹²⁷⁷ Considerant 13 de la PRIA.

¹²⁷⁸ COMISIÓN EUROPEA, *Propuesta de Reglamento del Parlamento europeo y del Consejo, por el que se establecen normas armonizadas en materia de inteligencia artificial...*, op. cit., 24 de abril de 2021, pàg. 15.

¹²⁷⁹ Considerant 27 de la PRIA.

¹²⁸⁰ Interessa referenciar aquí l'algoritme predictiu d'elaboració de perfils que emprava reconeixement facial de la Policia de Gal·les del Sud, que va ser considerat contrari a l'article 8 del CEDH, atès que no s'havien complert degudament tots els deures en relació amb la igualtat del sector públic (Public Sector

funcionament d'infraestructures essencials; educació i formació professional; treball, gestió de les persones treballadores i accés al treball per compte propi; accés i gaudi de serveis públics i privats essencials i els seus beneficis; assumptes relacionats amb l'aplicació de la llei¹²⁸¹; gestió de la migració, asil i control fronterer; i administració de justícia i altres processos democràtics¹²⁸². Aquests sistemes constitueixen una llista *numerus clausus*. Això és problemàtic, ja que poden existir nous àmbits en els quals es consideri que l'ús de la IA pot suposar riscos per a la salut, la seguretat i els drets fonamentals o inclús existeixen àmbits que ja els causen i no estan inclosos en l'annex III, com per exemple, els sistemes sanitaris, que no sempre seran emprats com a components de seguretat de productes inclosos en l'annex II¹²⁸³. Malgrat que l'article 7 de la PRIA permet a la Comissió incloure sistemes dins els àmbits ja especificats¹²⁸⁴, s'ha reclamat una clàusula més flexible de modificació i ampliació dels sistemes de l'annex III¹²⁸⁵.

El considerant 33 de la PRIA estableix que les imprecisions tècniques dels sistemes d'IA destinats a la identificació biomètrica remota de les persones físiques pot donar lloc a resultats esbiaixats i tenir conseqüències discriminatòries¹²⁸⁶. En relació amb sistemes d'IA de risc alt que intervenen en l'àmbit de l'educació¹²⁸⁷, el treball¹²⁸⁸ i l'accés a serveis

Equality Duties (PSED)), al no haver examinat si el sistema de reconeixement facial produïa discriminació indirecta per raó de sexe o origen ètnic. Vegeu: (Bridges) contra CCSWP and SSHD [2019] EWHC 2341 (Admin), par. 210. Disponible a: <https://www.judiciary.uk/wp-content/uploads/2020/08/R-Bridges-v-CC-South-Wales-ors-Judgment.pdf>. Darrera visita el 2 de setembre de 2023.

¹²⁸¹ El considerant 40 de la PRIA estableix directament que amb la finalitat d'evitar el risc de possibles biaixos, errors i opacitats, s'han de considerar de risc alt aquells sistemes d'IA l'objecte dels quals és ajudar a les autoritats judicials a investigar i interpretar els fets, el dret i a aplicar la llei a uns fets concrets.

¹²⁸² Annex III de la PRIA. En l'article 7 de la PRIA s'atorguen poders a la Comissió Europea per modificar la llista de l'annex III mitjançant l'addició de sistemes d'IA de risc alt quan el sistema d'IA comporti el risc de causar un perjudici a la salut i la seguretat, o el risc de tenir repercussions negatives en els drets fonamentals que sigui equivalent o major als riscos de perjudicis associats als sistemes de risc alt que ja es contemplen en l'annex III, tenint en compte certs criteris, entre ells, en l'apartat f, la mesura en què les persones que podrien patir tal perjudici o repercussions negatives es trobin en una posició de vulnerabilitat respecte l'usuari d'un sistema d'IA, en particular degut al desequilibri de poder o coneixement que ambdós posseeixen, les circumstàncies econòmiques o socials, o la seva edat (article 7.2.f).

¹²⁸³ SORIANO ARNANZ, A., "La propuesta de Reglamento..." *op. cit.*, 2021, pàg. 12.

¹²⁸⁴ Article 7 de la PRIA.

¹²⁸⁵ SORIANO ARNANZ, A., "La propuesta de Reglamento..." *op. cit.*, 2021, pàg. 12.

¹²⁸⁶ Considerant 33 de la PRIA.

¹²⁸⁷ Considerant 35 de la PRIA.

¹²⁸⁸ Considerant 36 de la PRIA. En el cas de l'àmbit del treball, la gestió de les persones treballadores i l'accés al treball per compte propi (sobretot en l'àmbit de la contractació i la selecció de personal), aquests sistemes poden perpetuar patrons històrics de discriminació, per exemple contra dones, certs grups d'edat, persones amb discapacitat, persones d'origens racials o ètnics concrets o amb una orientació sexual determinada, durant tot el procés de contractació i en l'avaluació, la promoció o la retenció de persones en relacions contractuals de tipus laboral.

públics i privats essencials¹²⁸⁹, s'estableix textualment que els sistemes d'IA emprats en aquest àmbit poden perpetuar patrons històrics de discriminació¹²⁹⁰. El considerant 38 també precisa que les actuacions de les autoritats encarregades de l'aplicació de la llei que impliquen determinats usos de sistemes d'IA, es caracteritzen per un important desequilibri de poder, que, si no estan entrenats amb dades de bona qualitat, no compleixen els requisits oportuns en termes de precisió i solidesa, o no es dissenyen i proven degudament abans d'introduir-se en el mercat o posar-se en servei, poden senyalar a persones de manera discriminatòria, incorrecta o injusta¹²⁹¹. També es reconeix que la IA emprada en el control policial i migratori opera en un context en el qual les persones es veuen en una posició de vulnerabilitat, on es important que el sistema d'IA no sigui discriminatori¹²⁹².

Malgrat que la PRIA reconeix els efectes discriminatoris que poden tenir tots aquests sistemes, no mitiga adequadament els danys potencials que poden patir grups oprimits i que per natura agreugen la discriminació i la desigualtat. És a dir, en no ignorar l'existència dels efectes discriminatoris i, tot i així, no prohibir l'ús dels sistemes d'IA que els poden causar, en devalua la importància. En comptes de prohibir aquests sistemes d'IA, s'estableixen uns requisits que han de complir, que en la majoria dels casos¹²⁹³, són autoavaluables pel mateix proveïdor que vol comercialitzar-los (com veurem en l'apartat 4 d'aquest capítol).

En el capítol II del Títol III de la PRIA, s'estableixen els requisits per aplicar als sistemes

¹²⁸⁹ En el cas de sistemes d'IA orientats a l'accés i el gaudi de determinats serveis i ajudes essencials de caràcter públic i privat necessaris per a que les persones participin en la societat o tinguin unes condicions de vida millors, poden discriminar a persones o grups o perpetuar patrons històrics de discriminació, per exemple, per motius d'origen racial o ètnic, discapacitat, edat, orientació social o generar noves formes d'efectes discriminatoris, com també afectar de manera considerable als mitjans de subsistència de les persones i poden infringir els seus drets fonamentals com el dret a la protecció social, no discriminació, dignitat humana o tutela judicial efectiva, per tant, s'han de considerar de risc alt.

¹²⁹⁰ Considerant 35, 36 i 37 de la PRIA.

¹²⁹¹ Considerant 38 de la PRIA. A més, s'estableix que pot impedir l'exercici de drets fonamentals processals com la tutela judicial efectiva, dret a un/a jutge/ssa imparcial, drets de defensa i presumpció d'innocència, sobretot quan els sistemes d'IA no són suficientment transparents i explicables ni estan ben documentats.

¹²⁹² Considerant 39 de la PRIA. Estableix aquest considerant que aquells sistemes d'IA emprats en la gestió de migració, asil i control fronterer afecten a persones que amb freqüència es troben en una situació especialment vulnerable i depenen del resultat de les autoritats públiques competents, essent important que els sistemes d'IA que s'emprin en aquests contextos siguin precisos, no discriminatoris i transparents, a fi de garantir que es respectin els drets fonamentals de les persones afectades i en particular els drets de lliure circulació, no-discriminació, intimitat personal, protecció de dades personals, protecció internacional i bona administració.

¹²⁹³ Només en el cas de la identificació biomètrica remota s'obliga a que en l'avaluació de conformitat hi participi un òrgan notificat.

d'IA de risc alt, que es refereixen a la qualitat dels conjunts de dades emprats¹²⁹⁴, la documentació tècnica¹²⁹⁵ i el registre¹²⁹⁶, la transparència i la comunicació d'informació a usuaris¹²⁹⁷, la vigilància humana¹²⁹⁸, i la solidesa, la precisió i la ciberseguretat¹²⁹⁹ (que abordem amb profunditat en l'apartat 3 d'aquest capítol).

En el moment en què s'acredita mitjançant una autoavaluació o avaluació per tercers, que el producte que conté un sistema d'IA de risc alt que es vol comercialitzar compleix els requisits, durà el certificat "CE" per acreditar la conformitat amb el Reglament i podrà circular lliurement pel mercat intern. Els estats membres no han de crear obstacles injustificats a la introducció en el mercat i a la posada en servei de sistemes d'IA de risc alt que compleixin els requisits establerts i duen el marcatge "CE" de conformitat¹³⁰⁰.

En el capítol III del Títol III de la PRIA, s'imposa un conjunt d'obligacions horitzontals a qui proveeix sistemes d'IA de risc alt. També s'estableixen obligacions per les persones usuàries i altres participants en la importació, distribució i representació autoritzada. En el capítol IV del Títol III, s'estableix un marc pel qual els organismes notificats (organismes d'avaluació de la conformitat designats d'acord amb el Reglament i la legislació d'harmonització pertinent de la UE¹³⁰¹) han de participar en els procediments de conformitat com a tercers independents i en el capítol V del Títol III, s'expliquen en

¹²⁹⁴ Article 10 de la PRIA.

¹²⁹⁵ Article 11 de la PRIA.

¹²⁹⁶ Article 12 de la PRIA.

¹²⁹⁷ Article 13 de la PRIA.

¹²⁹⁸ Article 14 de la PRIA.

¹²⁹⁹ Article 15 de la PRIA. El considerant 50 de la PRIA estableix que la solidesa tècnica és un requisit clau per als sistemes d'IA de risc alt, que han de tenir la capacitat de resiliència cap als riscos associats a les limitacions del sistema (per exemple, errors, incoherències o situacions inesperades), així com a accions malicioses que poden posar en perill la seva seguretat i donar lloc a conductes perjudicials o indesitjables per altres motius. La incapacitat de protegir les persones enfront aquests riscos podria tenir conseqüències per la seguretat o afectar de manera negativa als drets fonamentals, per exemple, degut a l'adopció de decisions equivocades o a que el sistema d'IA en qüestió generi una informació de sortida errònia i esbiaixada. L'article 15.3, en aquest sentit, afirma que els sistemes d'IA de risc alt seran resistent a errors i incoherències que puguin sortir en els propis sistemes o en l'entorn on operen, en particular a causa de la seva interacció amb persones físiques o altres sistemes. Des de la perspectiva de la discriminació algorítmica, i tenint en compte els biaixos existents en les bases de dades, els biaixos implícits i explícits de les persones que dissenyen, desenvolupen o posen en funcionament un sistema d'IA, o la discriminació existent en la societat, entre altres, aquesta afirmació sembla molt categòrica.

¹³⁰⁰ Considerant 67 de la PRIA.

¹³⁰¹ Article 3.22 de la PRIA.

detall els procediments d'avaluació de la conformitat que s'han de seguir per cada tipus de sistemes d'IA de risc alt ¹³⁰².

2.3. Sistemes de risc baix o mínim

Després d'establir normes específiques pels sistemes d'IA de risc alt, el Títol IV de la PRIA estableix normes que es poden aplicar tant a sistemes de risc alt com a sistemes que no s'integrin en aquest grup. L'article 52 de la norma estableix que s'aplicaran obligacions de transparència als sistemes que interactuïn amb éssers humans, s'utilitzin per detectar emocions o determinar l'associació a categories socials concretes a partir de dades biomètriques o que generin o manipulin contingut. Quan una persona interactuï amb un sistema d'IA o les seves emocions o característiques siguin reconegudes per mitjans automatitzats, es precis informar-la de tal circumstància ¹³⁰³. El Parlament Europeu en les esmenes a la proposta de la Comissió precisa que, quan sigui apropiat i rellevant, la informació haurà d'incloure quines funcions d'intel·ligència artificial estan activades, si hi ha supervisió humana, qui és responsable del procés de presa de decisió, així com els drets i processos existents que permeten a les persones oposar-se a que se'ls apliqui un d'aquests sistemes i a sol·licitar remei judicial contra decisions preses o danys ocasionats per aquests sistemes, incloent el seu dret a sol·licitar una explicació ¹³⁰⁴.

Segons el considerant 70, aquests sistemes d'IA poden comportar riscos específics de suplantació o falsificació, amb independència de si són classificats com de risc alt o no, per tant se'ls han d'aplicar obligacions de transparència específiques, sense perjudici dels requisits aplicables als sistemes de risc alt ¹³⁰⁵. Els usuaris, hauran de notificar a les persones físiques que estan interactuant amb un sistema d'IA -llevat que sigui evident que ho estan fent per les circumstàncies i el context del seu ús ¹³⁰⁶-, i informar que estan

¹³⁰² COMISIÓN EUROPEA, *Propuesta de Reglamento del Parlamento europeo y del Consejo, por el que se establecen normas armonizadas en materia de inteligencia artificial...*, *op. cit.*, 24 de abril de 2021, pàg. 16.

¹³⁰³ Article 52 de la PRIA.

¹³⁰⁴ EUROPEAN PARLIAMENT, *Amendments adopted by the European Parliament on 14 June 2023...*, *op. cit.*, 14 June 2023.

¹³⁰⁵ Considerant 70 de la PRIA.

¹³⁰⁶ Article 52.1 de la PRIA. No s'aplicarà si estan autoritzats per la llei amb l'objectiu de prevenir infraccions penals.

exposades a un sistema de reconeixement d'emocions o un sistema de categorització biomètrica (que suposa l'ús d'IA pel reconeixement de certs trets d'identitat com la raça, la identitat de gènere o la discapacitat)¹³⁰⁷. Aquesta informació i aquestes notificacions han de ser accessibles a persones amb discapacitat i a més, els usuaris que emprin un sistema d'IA per generar o manipular imatges, arxius d'àudio o vídeos que semblin versemblantment persones, llocs o successos reals i puguin induir erròniament una persona a pensar que són autèntics, han de comunicar que han estat creats o manipulats de manera artificial, etiquetant el contingut generat per la IA i indicant el seu origen artificial¹³⁰⁸.

La norma considera usuari a la persona física o jurídica, autoritat pública, agència o organisme d'un altre tipus sota l'autoritat de la qual s'utilitzi el sistema d'IA, llevat quan el seu ús s'emmarqui en una activitat personal de caràcter no professional (per tant es tracta de les persones que usen els sistemes d'IA professionalment i no les persones que veuen els seus drets afectats per un sistema d'IA¹³⁰⁹). Aquesta definició restringida d'usuari que crea l'exempció d'activitats personals no professionals, suposa que sigui un repte fer aquesta distinció en el món en línia, atesa l'alça de marques personals i màrqueting d'"influencers" on s'empren filtres generats per la IA (per exemple, Snapchat i TikTok)¹³¹⁰.

També, com ja hem esmentat, és una obligació de transparència que només informa, i no ajuda les persones a entendre, establir una posició crítica i desafiar sistemes d'IA perjudicials i opacs, sense posicionar aquests sistemes com a parts d'estructures de poder sociotècniques més àmplies, que creen efectes discriminatoris¹³¹¹. Greenleaf augura que molts sistemes d'IA cauran en la categoria de sistemes de risc baix, que no generen obligacions¹³¹². En el mateix sentit, Cotino Hueso afirma que des del nivell més bàsic dels sistemes automatitzats i algorítmics és més que possible que el sistema generi lesions

¹³⁰⁷ Article 52.2 de la PRIA. No s'aplicarà si estan autoritzats per la llei amb l'objectiu de prevenir infraccions penals.

¹³⁰⁸ Article 52.3 de la PRIA. Existeixen exempcions per prevenció del crim autoritzat o necessitat d'exercir una llibertat d'expressió o d'art i ciències.

¹³⁰⁹ Article 3.4 de la PRIA.

¹³¹⁰ VEALE, M. i ZUIDERVEEN, F., "Demystifying the Draft EU...", *op. cit.*, 2021, pàg. 108.

¹³¹¹ BERENDT, B., "The AI Act Proposal...", *op. cit.*, 2022, pàg. 45. L'autora també opina que la transparència, malgrat que estigui aparellada amb la qualitat de les dades, seguretat, supervisió humana i documentació, no és suficient per prevenir la discriminació.

¹³¹² GREENLEAF, G., *The "Brussels Effect" of the EU's "AI Act" on Data Privacy Outside Europe*, Privacy Laws and Business, International Report, núm. 171, Sydney, 2021, pàg. 5.

de drets o interessos i vulneracions de la legalitat¹³¹³.

A tall de recapitulació, si bé és cert que és positiu en relació amb la discriminació algorítmica que es prohibeixi en la PRIA l'ús de certs sistemes d'IA que porten a un tracte perjudicial o desfavorable de certes persones o grups, es fa en un context relativament estret. Quant als sistemes de risc alt, reconèixer tots els efectes discriminatoris que poden tenir, els danys potencials que per natura agreugen la discriminació i desigualtat que poden patir grups oprimits i no prohibir-los, en devalua la importància. L'establiment de certs requisits, que en la majoria dels casos són autoavaluables pel mateix proveïdor que vol comercialitzar-los, no sembla una protecció suficient. A més, molts sistemes d'IA podran considerar-se de risc baix, i no generaran obligacions, quan potser, aplicats a un context determinat tindran una incidència molt més negativa d'aquella esperada. Tanmateix, malgrat presentar una estructura que considerem insuficient per abordar la discriminació algorítmica, tot seguit, analitzem els elements de la PRIA que considerem que poden ser rellevants en la protecció envers la discriminació algorítmica.

3. Elements de la PRIA rellevants per la protecció envers la discriminació

Segons la PRIA, està garantida la coherència de la norma amb la CDFUE i el dret derivat de la UE vigent en matèria de protecció de dades, protecció dels consumidors, no discriminació i igualtat de gènere. Es complementa el dret derivat de la UE vigent en matèria de no discriminació en establir requisits específics que tenen per objecte reduir al mínim el risc de discriminació algorítmica en el cas de sistemes d'IA classificats com de risc alt, en particular en relació al disseny i la qualitat dels conjunts de dades emprats per desenvolupar sistemes d'IA, els quals van acompanyats d'obligacions referents a la realització de proves, la gestió de riscos, la documentació i registre, i la vigilància humana durant tot el cicle de vida dels sistemes¹³¹⁴.

¹³¹³ COTINO HUESO, L., “Los usos de la inteligencia artificial en el sector público...”, *op. cit.*, 2023, pàg. 228.

¹³¹⁴ COMISIÓN EUROPEA, *Propuesta de Reglamento del Parlamento europeo y del Consejo, por el que se establecen normas armonizadas en materia de inteligencia artificial...*, *op. cit.*, 24 de abril de 2021, pàg. 4.

Aquests requisits també es poden complir voluntàriament en el cas de sistemes que no es considerin com de risc alt però que igualment es comercialitzin. En relació amb la discriminació algorítmica, malgrat tots tenen rellevància, creiem que necessiten un abordatge més incisiu el disseny i la qualitat dels conjunts de dades (1), la vigilància humana durant tot el cicle de vida dels sistemes (2) i l'exigència de documentació i registre (3). També volem analitzar, per l'aportació que podrien fer a la protecció envers la discriminació algorítmica, la creació que estableix la PRIA d'una base de dades pública (4), i d'espais de prova (5).

3.1. Disseny i qualitat de les dades

El model d'adequació que presenta la PRIA per als sistemes d'IA de risc alt, focalitza la seva atenció en la qualitat de les dades¹³¹⁵. El considerant 44 estableix que molts sistemes d'IA necessiten dades de qualitat alta per funcionar correctament, en especial quan es duen a terme tècniques que impliquen l'entrenament de models (aprenentatge automàtic), amb l'objectiu de garantir que el sistema d'IA de risc alt funcioni de la manera prevista i en condicions de seguretat i no es converteixi en una font d'algun dels tipus de discriminació prohibits pel dret de la UE¹³¹⁶.

L'article 10.1 de la PRIA estableix certs criteris de qualitat per als conjunts de dades d'entrenament¹³¹⁷, validació¹³¹⁸ i prova¹³¹⁹ que s'emprin en els sistemes d'IA que utilitzin tècniques que impliquin entrenament dels models amb dades¹³²⁰. En el segon apartat de l'article s'estableix com a criteri la submissió de les dades a pràctiques adequades de

¹³¹⁵ ZUDDAS, P., "Decisión algorítmica y discriminación", *op. cit.*, 2022, pàg. 10.

¹³¹⁶ Considerant 44 de la PRIA.

¹³¹⁷ L'article 3.29 de la PRIA estableix que són dades d'entrenament aquelles dades usades per entrenar un sistema d'IA mitjançant l'ajust dels seus paràmetres que es poden entrenar, entre els quals s'inclouen els passos d'una xarxa neuronal.

¹³¹⁸ L'article 3.30 de la PRIA estableix que les dades de validació són dades emprades per proporcionar una avaluació del sistema d'IA entrenat i adaptar els seus paràmetres que no es poden entrenar i el seu procés d'aprenentatge, entre altres coses per evitar el sobre ajust. El conjunt de dades de validació pot ser un conjunt de dades independent o formar part del conjunt de dades d'entrenament, ja sigui com una divisió fixa o variable.

¹³¹⁹ L'article 3.31 de la PRIA estableix que les dades de prova són dades emprades per proporcionar una avaluació independent del sistema d'IA entrenat i validat, amb la finalitat de confirmar el funcionament previst del sistema abans de la seva comercialització o posada en servei.

¹³²⁰ Article 10.1 de la PRIA.

governança i gestió de dades que se centraran en l'elecció del disseny adequat; la recopilació de dades; les operacions oportunes de preparació de les dades (anotació, etiquetatge, depuració, enriquiment i agregació); la formulació dels supòsits pertinents, fonamentalment pel que fa a la informació que, atenent-s'hi, les dades mesuren i representen; l'avaluació prèvia de la disponibilitat, qualitat i adequació dels conjunts de dades necessaris; l'examen atenent a possibles biaixos; i la detecció de possibles llacunes o deficiències en les dades i la forma d'esmenar-les¹³²¹.

En el tercer apartat s'estableix com a criteri que les dades d'entrenament, validació i prova seran pertinents i representatives, no tindran errors i estaran completes, amb les propietats estadístiques adequades, també en allò que respecta a les persones o grups de persones en referència a les quals es pretengui emprar el sistema de risc alt¹³²², afegint-se en el quart apartat que tindran en compte, en la mesura necessària en funció de la seva finalitat prevista, les característiques o elements particulars del context geogràfic, conductual o funcional específic en el qual es pretengui emprar el sistema de risc alt¹³²³. Finalment, com ja hem esmentat anteriorment, en el cinquè apartat es permet, en la mesura en què sigui estrictament necessari per garantir la vigilància, detecció i correcció de biaixos associats a sistemes d'IA de risc alt, tractar categories especials de dades¹³²⁴.

A més, en el considerant 45 de la PRIA es fa referència explícita als espais comuns europeus de dades establerts per la Comissió i la facilitació de l'intercanvi de dades entre empreses i amb els governs en favor de l'interès públic, per brindar un accés fiable, responsable i no discriminatori a dades d'alta qualitat amb els quals entrenar, validar i provar els sistemes d'IA¹³²⁵.

L'article 10 de la PRIA no especifica com verificar que els criteris de qualitat de dades s'han assolit ni sembla viable que es puguin assolir tots. Apunten, en aquest sentit, Veale i Zuiderveen que, si bé és cert que alguns requeriments semblen difícils d'assolir, per exemple, l'exigència de l'article 10.3 de bases de dades lliures d'errors i completes -les

¹³²¹ Article 10.2 de la PRIA.

¹³²² Article 10.3 de la PRIA.

¹³²³ Article 10.4 de la PRIA.

¹³²⁴ Article 10.5 de la PRIA.

¹³²⁵ Considerant 45 de la PRIA. Per exemple en l'àmbit de la salut, l'espai europeu de dades sanitàries facilitarà l'accés no discriminatori a dades sanitàries i l'entrenament, a partir d'aquests conjunts de dades, d'algoritmes d'intel·ligència artificial d'una manera segura, oportuna, transparent i fiable que respecti la privacitat i contant amb la deguda governança institucional.

quals habitualment no ho són¹³²⁶, segons el considerant 44, les bases de dades només han d'assolir aquests requisits potencials de manera suficient i en vista de la finalitat del sistema¹³²⁷.

També, com apunta Zuddas, tant l'article 10 com l'article 17 (sistema de gestió de qualitat), semblen reflectir la consciència dels riscos associats a l'ús dels sistemes d'IA, però, davant els sistemes d'aprenentatge autònom, sembla molt complicat establir remeis preventius. En aquests casos, s'hauria de recórrer al principi de precaució, que requeriria en la seva vessant més estricta, no emprar la intel·ligència artificial quan el risc de distorsions potencialment generables no pugui ser quantificat, o en la vessant més flexible, introduir procediments que incorporin l'anàlisi de risc en la implementació del sistema informàtic. En aquest sentit, l'article 9 de la PRIA, preveu un sistema de gestió de riscos (per als sistemes classificats com de risc alt) continu, sistemàtic i constantment actualitzat que inclou accions preventives en la fase de planificació i accions posteriors de mitigació i control de risc i obligacions d'informació de l'usuari¹³²⁸.

La PRIA reconeix que els sistemes d'IA perpetuen patrons històrics de discriminació, però, en focalitzar la mitigació de la discriminació en la qualitat de les dades s'infravaloren les causes i l'abast de la discriminació. Si bé qui desenvolupa els sistemes d'IA podrà predir i prevenir alguns biaixos, moltes de les aplicacions previstes en l'annex III inevitablement exacerbaran desigualtats estructurals, ja que els sistemes d'IA es desenvolupen en un context més ampli de desigualtats estructurals. Confiant en mesures tècniques per contrarestar el biaix com a resposta a la discriminació, no només es reforça la convicció que suprimir el biaix d'aquests sistemes pot ser possible, sinó que s'imposa la perspectiva que el des-biaix tècnic és completament efectiu per suprimir la discriminació algorítmica. Com planteja Roig en relació amb l'ús de tecnologies garants en l'àmbit de la privadesa, el seu ús no sempre és sinònim de protecció efectiva de la privadesa, algunes només serveixen en certs contextos i no totes són igualment protectores¹³²⁹. Afegeix que la tecnologia garant inclús pot incorporar nous riscos, donar una falsa sensació de seguretat o ser usada per legitimar una mesura reguladora¹³³⁰.

¹³²⁶NORTHCUTT, C.G. *et al.*, "Pervasive Label Errors in Test Sets Destabilize Machine Learning Benchmarks", 2021, pàg. 1-24. Disponible a: <https://arxiv.org/> Darrera visita el 2 de setembre de 2023.

¹³²⁷VEALE, M. i ZUIDERVEEN, F., "Demystifying the Draft EU...", *op. cit.*, 2021, pàg. 103.

¹³²⁸ZUDDAS, P., "Decisión algorítmica y discriminación", *op. cit.*, 2022, pàg. 11.

¹³²⁹ROIG, A., "Límits del principi de privadesa pel disseny", *op. cit.*, 2022, pàg. 178-179.

¹³³⁰*Ídem*, pàg. 184.

Adicionalment, la idea del des-biaix tècnic i l'ús de les millors tècniques disponibles per des-esbiaixar els sistemes d'IA es focalitza en els errors tècnics del sistema, i desvia l'atenció de qüestionar-se la seva adequació inclús quan no existeixen errors, per exemple, en el casos de sistemes de vigilància, quan funcionen adequadament i sense errors i igualment produeixen efectes indesitjables i reproduïxen patrons de discriminació¹³³¹. De manera que, centrar-se en la correcció dels errors tècnics és important sempre que no es cregui que és suficient per brindar equitat en els sistemes d'IA.

3.2. Documentació tècnica per dur a terme un plet de discriminació

Segons el considerant 46 de la PRIA, disposar d'informació sobre la manera en què s'ha desenvolupat el sistema d'IA i sobre el seu funcionament durant tot el cicle de vida és essencial per verificar si es compleixen els requisits presents al Reglament¹³³². Per a tal finalitat, es precis tenir registres¹³³³ i disposar de documentació tècnica que contingui la informació necessària per avaluar si el sistema en qüestió compleix els requisits pertinents¹³³⁴. No obstant això, cal puntualitzar que perquè aquest requisit participi a millorar la fiscalització dels sistemes d'IA de risc alt, resulta necessari que el compliment no sigui merament formal¹³³⁵.

L'article 11 de la PRIA estableix que qui proveeixi un sistema d'IA de risc alt haurà de preparar documentació tècnica abans la seva introducció en el mercat o posada en servei i la mantindrà actualitzada¹³³⁶. Aquesta documentació es redactarà de manera que demostrï que el sistema d'IA compleix amb els requisits establerts i proporcionarà a les

¹³³¹ KATELL, M., YOUNG, M., DAILEY, D., HERMAN, B., GUETLER, V., TAM, A., BINZ, C., RAZ, D., i KRAFFT, P., "Toward situated interventions for algorithmic equity...", *op. cit.*, 2020, pàg. 52.

¹³³² Considerant 46 de la PRIA.

¹³³³ Article 12 de la PRIA. En aquest article es precisa que els sistemes d'IA de risc alt es dissenyaran i desenvoluparan amb capacitat per generar arxius de registre mentre estiguin en funcionament per garantir la traçabilitat del funcionament del sistema durant el seu cicle de vida, incloent, el període de cada ús del sistema, la base de dades de referència amb la qual el sistema ha confrontat les bases de dades, les dades d'entrada amb els quals la recerca ha establert una correspondència i la identificació de persones físiques implicades en la verificació dels resultats.

¹³³⁴ Considerant 46 de la PRIA.

¹³³⁵ SORIANO ARNANZ, A., "La propuesta de Reglamento..." *op. cit.*, 2021, pàg. 17.

¹³³⁶ Article 11.1 de la PRIA.

autoritats nacionals competents i als organismes notificats la informació que necessitin per avaluar que el sistema d'IA compleix els requisits¹³³⁷. L'annex IV de la norma estableix el contingut que com a mínim haurà de contenir la documentació tècnica, i entre altres requisits, aquells pertinents en relació amb la discriminació són:

- 2.b. Una descripció detallada de les especificacions del disseny del sistema, a saber, la lògica general del sistema d'IA i dels algoritmes¹³³⁸; les opcions clau de disseny, en particular, la justificació lògica i les hipòtesis plantejades, també respecte a les persones i grups de persones en relació amb les quals es pretén emprar el sistema, les principals opcions de classificació; allò que el sistema està dissenyat per optimitzar i la pertinència dels diferents paràmetres; les decisions adoptades sobre qualsevol possible efecte de compensació respecte a les solucions tècniques adoptades per donar compliment als requisits¹³³⁹.
- 2.d. Una descripció detallada dels requisits sobre dades en forma de fitxes tècniques que descriu les metodologies i tècniques d'entrenament emprats, inclosa la informació sobre la procedència dels conjunts de dades, els seu abast i les seves característiques principals, com es van obtenir i seleccionar les dades, els procediments d'etiquetatge (per exemple, per l'aprenentatge supervisat), les metodologies de depuració de dades (per exemple, la detecció d'anomalies)¹³⁴⁰.
- 2.g. Una descripció detallada dels procediments de validació i prova emprats, inclosa la informació sobre les dades de validació i prova emprades i les seves característiques principals, els paràmetres emprats per mesurar la precisió, la solidesa, la ciberseguretat i el compliment d'altres requisits, *així com els efectes potencialment discriminatoris*¹³⁴¹.
- 3. Informació detallada sobre el seguiment, funcionament i control del sistema d'IA, en particular respecte a les seves capacitats i limitacions de funcionament, inclosos els nivells de precisió per a les persones o grups de persones específics

¹³³⁷ Article 11.2 de la PRIA.

¹³³⁸ El Parlament Europeu en les seves esmenes al text plantejat per la Comissió canvia la frase per la següent: una descripció de l'arquitectura, les especificacions del disseny del sistema, algoritmes i estructures de dades incloent la descomposició dels seus components i interfícies, com es relacionen entre sí i com proporcionen el tractament global o la lògica del sistema d'IA (...).

¹³³⁹ Apartat 2.b de l'annex IV de la PRIA.

¹³⁴⁰ Apartat 2.d de l'annex IV de la PRIA.

¹³⁴¹ Apartat 2.g de l'annex IV de la PRIA.

respecte als que es preveu emprar el sistema i els nivell de precisió general esperat en relació amb la finalitat prevista, els resultats no desitjats previsibles i les fonts de risc per a la salut i la seguretat, els drets fonamentals i la discriminació tenint en compte la finalitat prevista del sistema d'IA¹³⁴².

El considerant 47 de la PRIA estableix que s'ha d'exigir cert grau de transparència respecte els sistemes d'IA de risc alt per esmenar l'opacitat que pot fer incomprendibles alguns dels sistemes o massa complexos per a les persones físiques. Els usuaris han de ser capaços d'interpretar la informació de sortida del sistema i d'emprar-la adequadament. En conseqüència, els sistemes d'IA de risc alt han d'anar acompanyats de la documentació i les instruccions d'ús oportunes i incloure informació clara i concisa, en particular sobre els possibles riscos pels drets fonamentals i de discriminació quan correspongui¹³⁴³.

Les autoritats de vigilància i supervisió existents estaran facultades per sol·licitar qualsevol documentació que es conservi en virtut del Reglament i accedir a ella¹³⁴⁴. També, i més rellevant en el cas de supòsits que incloguin discriminació, segons l'article 64.3 i el considerant 79 de la PRIA, quan sigui necessari per al compliment del mandat, les autoritats o els organismes públics nacionals que supervisen l'aplicació del dret de la UE que protegeix els drets fonamentals, inclosos els organismes d'igualtat, també han de tenir accés a la documentació que es creï en virtut del Reglament¹³⁴⁵.

Qui proveeix el sistema d'IA no ha de publicar la documentació tècnica o donar-la excepte a organitzacions involucrades en la regulació o l'avaluació de conformitat. Diferents provisions, així i tot, indiquen que la informació ha de ser proveïda com una forma de transparència per les persones usuàries i registrada en la base de dades pública¹³⁴⁶. Això suposa que existirà documentació tècnica on s'explicitaran els possibles efectes discriminatoris que pot tenir un sistema d'IA de risc alt. Aquesta informació pot ser de suma rellevància en un plet de discriminació, atès que exigeix informació detallada sobre l'actuació del sistema d'IA en persones o grups de persones i inclús els efectes

¹³⁴² Apartat 3 de l'annex IV de la PRIA.

¹³⁴³ Considerant 47 de la PRIA.

¹³⁴⁴ COMISIÓN EUROPEA, *Propuesta de Reglamento del Parlamento europeo y del Consejo, por el que se establecen normas armonizadas en materia de inteligencia artificial...*, *op. cit.*, 24 de abril de 2021, pàg. 18.

¹³⁴⁵ Article 64.3 i considerant 79 de la PRIA.

¹³⁴⁶ VEALE, M. i ZUIDERVEEN, F., "Demystifying the Draft EU...", *op. cit.*, 2021, pàg. 103.

potencialment discriminatoris.

3.3. Supervisió humana dels sistemes de risc alt

L'article 14 de la PRIA requereix que els sistemes d'IA de risc alt es dissenyin i es desenvolupin de manera que puguin ser vigilats de manera efectiva per persones físiques durant el període que estiguin en ús, essent l'objectiu d'aquesta vigilància, prevenir o reduir al mínim els risc per a la salut, la seguretat o els drets fonamentals¹³⁴⁷.

L'article 14.3 de la PRIA, estipula que la persona proveïdora del sistema d'IA establirà mesures per a la vigilància humana del sistema, que permetran que a les persones que se'ls encomani la vigilància (art. 14.4.b) siguin conscients de la possible tendència a confiar automàticament o en excés en la informació de sortida generada per un sistema d'IA de risc alt (biaix d'automatització), sense estar massa clar com aquesta consciència serà suficient per evitar-lo¹³⁴⁸. També estableix que s'hauran d'establir mesures que permetin a les persones a les quals se'ls hi encomani la vigilància humana, decidir, en qualsevol situació concreta, no emprar el sistema d'IA de risc alt o desestimar, invalidar o revertir la informació de sortida que generi¹³⁴⁹. En el cas que es generessin resultats discriminatoris, la persona que dugués a terme la vigilància humana hauria de tenir mitjans per invalidar el resultat. Així i tot, el precepte dona llibertat als proveïdors per decidir la forma en la qual s'articularà tot el mecanisme, sense establir què es considera suficient intervenció humana ni fixar cap mandat específic¹³⁵⁰.

La visió de la supervisió humana durant tota la vida del sistema, s'integra en la perspectiva del sistema de gestió de riscos que s'exigeix com a requisit pels sistemes de risc alt en l'article 9 de la norma. En aquest article s'especifica que consistirà en un procés repetitiu continu que es durà a terme durant tot el cicle de vida del sistema, que requerirà actualitzacions sistemàtiques periòdiques, i constarà d'una etapa d'identificació i anàlisi dels riscos coneguts i previsibles de cada sistema; d'una etapa

¹³⁴⁷ Article 14 de la PRIA.

¹³⁴⁸ BERENDT, B., "The AI Act Proposal...", *op. cit.*, 2022, pàg. 43.

¹³⁴⁹ Article 14.4.d de la PRIA.

¹³⁵⁰ SORIANO ARNANZ, A., "La propuesta de Reglamento..." *op. cit.*, 2021, pàg. 19.

d'estimació i avaluació dels riscos que podran sorgir quan el sistema d'IA de risc alt en qüestió s'utilitzi conforme a la seva finalitat prevista i quan se li doni un ús indegut raonablement previsible; d'una etapa d'avaluació d'altres riscos que podrien sorgir a partir de l'anàlisi de dades recollits en el sistema de seguiment posterior a la comercialització¹³⁵¹; i una etapa d'adopció de les mesures oportunes de gestió de riscos¹³⁵².

A més, segons l'article 12 de la PRIA, es precisa també que els sistemes d'IA de risc alt es dissenyaran i desenvoluparan amb capacitat per generar arxius de registre mentre estiguin en funcionament per garantir la traçabilitat del funcionament dels sistema durant el seu cicle de vida, incloent en aquests arxius, la identificació de persones físiques implicades en la verificació dels resultats¹³⁵³.

En la protecció de dades, la supervisió humana típicament té relació amb la dignitat, en canvi, en la PRIA la supervisió humana té relació amb la minimització dels riscos per a la sanitat, la seguretat i els drets fonamentals (art. 14.2)¹³⁵⁴.

Segons Green, en un estudi fet en relació amb algoritmes emprats en el sector públic, la supervisió humana presenta dues mancances significatives, en primer lloc, que les persones no poden proveir de manera adequada la supervisió prevista i en segon lloc, l'assumpció incorrecta que una supervisió humana efectiva legitima l'ús d'algoritmes defectuosos i inexplicables¹³⁵⁵. És una figura que en lloc de proporcionar els beneficis dels algoritmes sense els riscos associats, justifica l'ús inadequat d'algoritmes i dificulta la rendició de comptes de les persones responsables¹³⁵⁶. Abans de desenvolupar un algoritme amb supervisió humana, apunta, qui el proveeix ha de dur a terme avaluacions preliminars sobre si les persones poden de manera efectiva supervisar l'algoritme i considerar si es apropiat usar un algoritme en un context determinat, tenint en compte que

¹³⁵¹ L'article 61 de la PRIA aborda el seguiment posterior a la comercialització i estableix que es basarà en un pla de seguiment posterior a la comercialització que formarà part de la documentació tècnica requerida pels sistemes de risc alt.

¹³⁵² Article 9.2 de la PRIA.

¹³⁵³ Article 12 de la PRIA. L'article estableix que a part de la identificació de persones físiques implicades en la verificació del resultat, també s'inclouran en els arxius de registre el període de cada ús del sistema, la base de dades de referència amb la qual el sistema ha confrontat les bases de dades, i les dades d'entrada amb els quals la recerca ha establert una correspondència.

¹³⁵⁴ VEALE, M. i ZUIDERVEEN, F., "Demystifying the Draft EU...", *op. cit.*, 2021, pàg. 103.

¹³⁵⁵ GREEN, B., "The flaws of policies requiring human oversight of government algorithms", *Computer Law and Security Review*, vol. 45, 2022, pàg. 16.

¹³⁵⁶ *Ídem, ibídem.*

els algoritmes no són efectius resolent qüestions definides pobrament¹³⁵⁷. En teoria, l'article 14 de la PRIA resol aquesta qüestió, atès que requereix que la vigilància sigui de manera efectiva.

3.4. Base de dades pública

L'article 60 de la PRIA estableix que la Comissió, amb la col·laboració dels estats membres, crearà i mantindrà una base de dades de la UE dels sistemes de risc alt que encaixin amb els paràmetres establerts en l'annex III¹³⁵⁸, que d'acord amb l'article 51 de la norma s'hauran de registrar en la base de dades abans de la seva introducció en el mercat o posada en servei¹³⁵⁹. Els proveïdors introduiran en la base de dades de la UE, entre altra informació, el nom, direcció i dades de contacte, el nom comercial del sistema d'IA i tota referència inequívoca addicional que permeti la seva identificació i traçabilitat, com també una descripció de la finalitat prevista per sistema¹³⁶⁰. Aquesta informació serà accessible per al públic¹³⁶¹. El Parlament Europeu en les esmenes proposades al text de la Comissió planteja que certs usuaris també hagin de registrar l'ús de sistemes d'IA en la base de dades¹³⁶².

La informació que detalla l'Annex VIII de la PRIA¹³⁶³, que s'ha d'introduir en la base de dades, sembla més aviat tècnica que no qualitativa. Així i tot, en l'exposició de motius de la PRIA s'explicita que per alimentar aquesta base de dades, els proveïdors d'IA estan obligats a facilitar informació significativa sobre els seus sistemes i l'avaluació de

¹³⁵⁷ *Ídem*, pàg. 15.

¹³⁵⁸ Article 60.1 de la PRIA.

¹³⁵⁹ Article 51 de la PRIA.

¹³⁶⁰ Tota la informació que s'ha d'introduir a la base de dades està detallada a l'Annex VIII de la PRIA. Interessa que el Parlament Europeu en les esmenes al text proposat per la Comissió afegeix que entre la informació a introduir en la base de dades s'haurà d'incorporar una descripció simple i comprensible de la finalitat del sistema d'IA, els components i les funcions basades en IA i una explicació bàsica de la lògica del sistema d'IA (punt 5). També, on sigui aplicable, les categories i la natura de les dades que probablement o que es preveu que seran tractades pel sistema d'IA (punt 5.a).

¹³⁶¹ Article 60.3 de la PRIA.

¹³⁶² EUROPEAN PARLIAMENT, *Amendments adopted by the European Parliament on 14 June 2023...*, *op. cit.*, 14 June 2023.

¹³⁶³ El Parlament Europeu en les seves esmenes al text de la Comissió divideix la informació detallada en l'Annex VIII de la PRIA en una Secció A (informació que han d'introduir els proveïdors) i Secció B (informació que han d'introduir els usuaris).

conformitat a la qual es sotmeten¹³⁶⁴. Malgrat que l'expressió “*informació significativa sobre els seus sistemes*” és imprecisa, suposa que existirà més informació accessible de la que hi havia fins ara.

Això significa que una persona que ha estat rebutjada per un sistema algorítmic de contractació, per exemple, podrà buscar exactament mitjançant quin sistema d'IA (atès que els sistemes d'IA es consideren de risc alt quan entren en l'àmbit del treball) s'ha dut a terme el procés de selecció, un primer pas per discutir, si es dona el cas, posteriorment, si la precisió dels resultats d'aquest sistema concret en certs grups de població es suficientment acurada, i si pot generar en el cas de no ser-ho, resultats discriminatoris.

L'objectiu de la base de dades, segons l'exposició de motius de la PRIA, és facilitar el seguiment de la Comissió i autoritats nacionals dels sistemes d'IA de risc alt independents amb implicacions principalment pels drets fonamentals¹³⁶⁵, com també permetre que altres persones interessades verifiquin si un sistema d'IA de risc alt compleix amb els requisits estipulats en el Reglament i exerceixin una vigilància reforçada a aquells sistemes d'IA que generin un risc alt pels drets fonamentals¹³⁶⁶. Segons Veale i Zuiderveen, podrà ajudar també a altres persones, presumiblement la societat civil o periodistes, a destapar sistemes il·lícits de IA¹³⁶⁷. Apunten igualment que en allò relatiu a persones usuàries finals afectades, si no tenen drets de reclamació clars, l'efectivitat i aplicació d'aquesta base de dades podrà ser significativament obstaculitzada¹³⁶⁸.

3.5. Espais de prova

L'article 53 de la PRIA disposa la creació d'espais de prova controlats establerts per les autoritats competents d'un o diversos estats membres o pel Supervisor Europeu de Protecció de dades, per proporcionar un entorn controlat que faciliti el desenvolupament, prova i validació de sistemes innovadors d'IA durant un període limitat abans de la seva

¹³⁶⁴ COMISIÓN EUROPEA, *Propuesta de Reglamento del Parlamento europeo y del Consejo, por el que se establecen normas armonizadas en materia de inteligencia artificial...*, *op. cit.*, 24 de abril de 2021, pàg. 14.

¹³⁶⁵ *Ídem*, pàg. 17.

¹³⁶⁶ *Ídem*, pàg. 14.

¹³⁶⁷ VEALE, M. i ZUIDERVEEN, F., “Demystifying the Draft EU...”, *op. cit.*, 2021, pàg. 112.

¹³⁶⁸ *Ídem, ibídem*.

comercialització o posada en servei¹³⁶⁹. Aquests espais de prova, a Espanya, ja s'empren, per exemple, en el sector financer. El títol II de la Ley 7/2020, de 13 de noviembre, para la transformación digital del sistema financiero, regula l'espai controlat de proves, que a escala europea i internacional s'anomena "regulatory sandbox", i engloba un conjunt de disposicions que emparen la realització controlada i delimitada de proves dins un projecte que pot aportar una innovació financera de base tecnològica aplicable al sistema financer, definida com aquella que pugui donar lloc a nous models de negoci, aplicacions, processos o productes amb incidència en el mercat financer¹³⁷⁰. També en el sector elèctric, es va aprovar el juliol de 2022 un Reial Decret amb el propòsit d'establir el marc general del banc de proves regulador en el sector elèctric¹³⁷¹.

Els estats membres hauran de vetllar perquè en la mesura en què els sistemes impliquin tractament de dades personals o s'inclouin en l'àmbit de supervisió d'altres autoritats nacionals, aquestes estiguin lligades al funcionament dels espais de prova controlats¹³⁷². Això significa que les autoritats nacionals d'igualtat podran estar implicades en el funcionament dels espais de prova controlats (vegeu apartat 2.3.2 del capítol VIII). Els bancs de proves interessen com a eines de coregulació amb les quals les autoritats nacionals poden evitar riscos quan es puguin detectar abans de la comercialització d'un producte.

Qualsevol risc significatiu per a la salut, la seguretat i els drets fonamentals detectat durant el procés de desenvolupament implicarà la mitigació immediata o, si no és possible, la suspensió del procés fins que es produeixi aquesta mitigació¹³⁷³.

Els espais de prova estan orientats com a mesures addicionals per impulsar la innovació¹³⁷⁴, per facilitar el desenvolupament i la prova de sistemes d'IA innovadors,

¹³⁶⁹ Article 53.1 de la PRIA.

¹³⁷⁰ Ley 7/2020, de 13 de noviembre, para la transformación digital del sistema financiero. BOE, núm. 300, de 14 de noviembre de 2020. España.

¹³⁷¹ Real Decreto 568/2022, de 11 de julio, por el que se establece el marco general del banco de pruebas regulatorio para el fomento de la investigación y la innovación en el sector eléctrico. BOE, núm. 166, de 12 de julio de 2022. España. En l'article 2.2 del Reial Decret es defineix el banc de proves regulador com un entorn experimental en el qual poden realitzar-se proves o assajos dins el marc d'execució de projectes pilot de manera segura i controlada, amb la finalitat de facilitar la investigació i la innovació en l'àmbit del sector elèctric.

¹³⁷² Article 53.2 de la PRIA.

¹³⁷³ Article 53.3 de la PRIA.

¹³⁷⁴ COMISIÓN EUROPEA, *Propuesta de Reglamento del Parlamento europeo y del Consejo, por el que se establecen normas armonizadas en materia de inteligencia artificial...*, op. cit., 24 de abril de 2021, pàg. 4.

sota una estricta vigilància reguladora abans de la seva introducció en el mercat o posada en funcionament¹³⁷⁵. Malgrat que estiguin creats per potenciar la innovació, donen una oportunitat per identificar les vulnerabilitats dels sistema i els efectes potencials que pot tenir en persones, organitzacions o en l'àmbit econòmic, social o ecològic; considerar l'abast de les dades i els possibles casos extrems en els quals el sistema podria fallar; i demostrar el rendiment del sistema en diferents conjunts de dades, incloent conjunts de dades externs que no s'hagin usat en la fase d'entrenament¹³⁷⁶.

En resum, la qualitat de les dades que s'instaura en la PRIA és ambiciosa però s'ha criticat que potser és impossible de complir (p. e. bases de dades lliures d'errors i completes). A més, instaura un discurs molt centrat en l'exhaustivitat de la correcció d'errors tècnics per abordar la discriminació, quan l'equitat en els sistemes d'IA depèn també d'altres factors. Quant a la documentació tècnica, és la primera norma que exigeix informació de pes relativa als efectes discriminatoris del sistema algorítmic i que potencialment pot ser útil en plets de discriminació. Tanmateix, l'accés a aquesta informació està restringit a autoritats de vigilància i supervisió i autoritats o organismes públics nacionals que supervisen l'aplicació del dret de la UE, que protegeix els drets fonamentals, inclosos els organismes d'igualtat, fet que limita la seva incidència. En relació amb la supervisió humana, no obstant que s'estableix que serà un mecanisme repetitiu continu que es realitzarà durant tot el cicle de vida del sistema, generarà registres i es durà a terme de manera efectiva, per l'experiència prèvia, segueix generant suspicàcies en relació amb la possibilitat que es creï una falsa sensació d'empara sense proveir realment de protecció legal a les persones i col·lectius. Quant a la creació i manteniment d'una base de dades de la UE dels sistemes de risc alt descrits en l'annex III de la PRIA, sembla un bona proposta i que pot tenir un impacte positiu en la detecció i desafiament de la discriminació algorítmica, malgrat que hauria de ser obligatori el registre de tots els sistemes d'IA i no només aquells de risc alt. Finalment, els espais de prova, malgrat que estiguin creats per potenciar la innovació, donen una oportunitat per identificar les vulnerabilitats dels sistema i obren possibilitats en la lluita antidiscriminació tenint en compte que els organismes d'igualtat nacionals poden estar implicats en el seu funcionament.

¹³⁷⁵ Considerant 71 de la PRIA.

¹³⁷⁶ HAMON, R, JUNKLEWITZ, H i SANCHEZ, I., *Robustness and Explainability of Artificial Intelligence*, JRC Technical Report, EU Science Hub, European Commission, Luxembourg, 2020, pàg. 23.

4. No discriminació i estandardització de sistemes d'IA

En aquest apartat estudiem el procés pel qual passarà el sistema d'IA abans i després de la seva comercialització i les implicacions d'aquest procés per controlar que els sistemes d'IA de risc alt comercialitzats siguin el menys discriminatoris possible.

La PRIA estableix en la seva regulació uns requisits essencials que es concreten mitjançant normes harmonitzades (que també anomenarem estàndards harmonitzats), creades per les organitzacions europees de normalització. Serà en els estàndards harmonitzats on s'hauran de traslladar les exigències concretes que hauran de complir els sistemes d'IA de risc alt en relació amb el principi de no-discriminació, fet que pot generar certes problemàtiques (4.1). Els proveïdors d'un sistema de risc alt, per garantir que aquest compleix els requisits essencials podran o bé passar una autoavaluació de conformitat, o bé una avaluació de conformitat a través d'un organisme notificat (4.2) per aconseguir el marcatge "CE" de conformitat, que els permetrà circular lliurement pel mercat intern de la UE (4.3). Una vegada comercialitzat, el proveïdor haurà d'establir un sistema de supervisió, en el qual les autoritats de vigilància de mercat (autoritats nacionals que s'encarreguen del control de productes que entren en el mercat de la UE, però no de manera específica de productes d'IA), tindran un paper principal (4.4).

4.1. Trasllet del principi de no-discriminació als estàndards creats per les organitzacions europees d'estandardització

En aquest apartat analitzem les dificultats del trasllat del principi de no-discriminació a les normes harmonitzades creades per les organitzacions europees d'estandardització. En primer lloc estudiem l'esquema del Nou Marc Legislatiu, pel qual es regeix la comercialització i posada en servei de sistemes d'IA de risc alt (4.1.1), per abordar en segon lloc, si aquesta sistemàtica, que propicia una delegació de poders legislatius a organitzacions privades (atès que les organitzacions europees de normalització són

organismes de dret privat), pot captar els matisos que suposa traslladar el principi de no-discriminació, en tant que dret fonamental, a normes harmonitzades (4.1.2).

4.1.1. Nou Marc Legislatiu i organitzacions Europees d'Estandardització

L'article 40 de de la PRIA estableix que els sistemes d'IA de risc alt que siguin conformes amb normes harmonitzades, o part d'aquestes, les referències de les quals s'hagin publicat en el Diari Oficial de la Unió Europea, seran conformes amb els requisits imposats pel Reglament¹³⁷⁷. El considerant 61 de la PRIA, estableix que els proveïdors han de complir les normes harmonitzades definides en el Reglamento (UE) 1025/2012 del Parlamento Europeo y del Consejo, de 25 de octubre de 2012, sobre la normalización europea¹³⁷⁸ (d'ara en endavant, Reglament 1025/2012), per demostrar la seva conformitat amb els requisits previstos en el Reglament¹³⁷⁹.

El Reglament 1025/2012 defineix les normes com una especificació tècnica adoptada per un organisme de normalització reconegut, d'aplicació repetida i continua, l'observança de la qual no és obligatòria. Una norma pot ser internacional, si l'adopta un organisme internacional de normalització; europea, si l'adopta una organització europea de normalització; harmonitzada¹³⁸⁰, si es tracta d'una norma europea adoptada arran d'una petició de la Comissió per l'aplicació de la legislació d'harmonització de la UE; o nacional, si es adoptada per un organisme nacional de normalització¹³⁸¹. Les normes

¹³⁷⁷ Article 40 de la PRIA. En l'article 41 s'especifica que quan no existeixin normes harmonitzades, la Comissió consideri que les normes harmonitzades pertinents són insuficient o que es necessari abordar qüestions específiques relacionades amb la seguretat o els drets fonamentals, la Comissió podrà adoptar, mitjançant actes d'execució, especificacions comunes.

¹³⁷⁸ Reglamento (UE) 1025/2012 del Parlamento Europeo y del Consejo, de 25 de octubre de 2012, sobre la normalización europea, por el que se modifican las Directivas 89/686/CEE y 93/15/CEE del Consejo y las Directivas 94/9/CE, 94/25/CE, 95/16/CE, 97/23/CE, 98/34/CE, 2004/22/CE, 2007/23/CE, 2009/23/CE y 2009/105/CE del Parlamento Europeo y del Consejo y por el que se deroga la Decisión 87/95/CEE del Consejo y la Decisión no 1673/2006/CE del Parlamento Europeo y del Consejo. DOUE, L 316/12, 14.11.2012.

¹³⁷⁹ Considerant 61 de la PRIA. En aquest considerant també s'estableix que les normes harmonitzades hauran de jugar un rol clau per proveir solucions tècniques als proveïdors per assegurar la conformitat amb el Reglament dels sistemes d'IA.

¹³⁸⁰ L'article 3.27 de la PRIA la defineix com una norma europea conforme a la definició que figura en l'article 2.1.c del Reglament 1025/2012.

¹³⁸¹ Article 2.1 del Reglament 1025/2012.

europnees, d'acord amb el considerant 4 del Reglament 1025/2012¹³⁸², són adoptades per les organitzacions europees de normalització¹³⁸³, a saber, el Comitè Europeu de Normalització (CEN)¹³⁸⁴, el Comitè Europeu de Normalització Electro-tècnica (Cenelec)¹³⁸⁵ i l'Institut Europeu de Normes de Telecomunicacions (ETSI)¹³⁸⁶. L'article 40 de la PRIA, estableix que els sistemes d'IA que siguin conformes amb normes harmonitzades són conformes als requisits prevists al Reglament. Aquestes normes harmonitzades, doncs, són normes europees (creades per organitzacions europees de normalització), adoptades arran d'una petició de la Comissió.

Aquest esquema regulador s'adapta al Nou Marc Legislatiu¹³⁸⁷ (una perspectiva comuna a la UE per regular certs productes com ascensors, dispositius mèdics, equipaments de protecció personal i jocs¹³⁸⁸), que es caracteritza per regulacions de seguretat del producte que especifiquen només els requisits essencials que han de complir els productes introduïts en el mercat intern de la UE per gaudir de llibertat de moviment, mentre que la tasca de donar a aquests requisits una forma més concreta es delega en les Organitzacions europees de normalització¹³⁸⁹. Per exemple, la PRIA estableix que les dades de validació i prova han de ser rellevants, representatives, sense errors i completes¹³⁹⁰ per assegurar que el sistema d'IA no esdevé una font de discriminació prohibida en la UE¹³⁹¹, sense indicar quines formes de biaixos estan prohibides en el marc legal existent i com el biaix algorítmic hauria de ser mitigat. El mateix s'aplica a l'article 13.1 de la PRIA referent a com els sistemes d'IA de risc alt han de ser dissenyats i desenvolupats de manera que s'asseguri que operen de manera suficientment transparent com per permetre a les persones usuàries interpretar de manera apropiada el resultat del sistema¹³⁹², deixant oberta la determinació de quin tipus de grau de transparència hauria de ser vist com a apropiat¹³⁹³.

L'esquema del Nou Marc Legislatiu és el següent: s'estableixen certs requisits (que ha

¹³⁸² Considerant 4 del Reglament 1025/2012.

¹³⁸³ En anglès, Standard Developing Organisations (SDOs).

¹³⁸⁴ En anglès, European Committee for Standardisation (CEN).

¹³⁸⁵ En anglès, European Committee for Electrotechnical Standardization (Cenelec).

¹³⁸⁶ En anglès, European Telecommunications Standards Institute (ETSI).

¹³⁸⁷ En anglès, New Legislative Framework.

¹³⁸⁸ VEALE, M. i ZUIDERVEEN, F., "Demystifying the Draft EU...", *op. cit.*, 2021, pàg. 102.

¹³⁸⁹ EBERS, M., "Standardizing AI. The Case...", *op. cit.*, 2022, pàg. 334.

¹³⁹⁰ Article 10.3 de la PRIA.

¹³⁹¹ Considerant 44 de la PRIA.

¹³⁹² Article 13.1 de la PRIA.

¹³⁹³ EBERS, M., "Standardizing AI. The Case...", *op. cit.*, 2022, pàg. 336.

d'acomplir un producte abans la seva posada en el mercat intern de la UE) en una regulació. Aquests requisits es desenvolupen en normes harmonitzades per part d'organitzacions europees de normalització, atès que la Comissió ho ha demanat per a l'aplicació de la legislació d'harmonització de la UE. El fabricant (en la terminologia emprada en la PRIA, proveïdor) ha de realitzar controls previs a la posada del producte en el mercat perquè aquest compleixi amb les normes harmonitzades. El fabricant, atès que ha detallat el disseny i procés de producció del producte, també és l'encarregat de completar l'avaluació de conformitat, amb l'objectiu d'obtenir el marcatge "CE", que dona la capacitat de circular lliurement pel mercat intern de la UE¹³⁹⁴. Que recaigui en el fabricant únicament el compliment de l'avaluació de conformitat distingeix el règim del Nou Marc Legislatiu de la regulació farmacèutica, on una autoritat pública (p. e. l'Agència Europea del Medicament) duu a terme avaluacions abans de rebre una aprovació per introduir el medicament al mercat¹³⁹⁵.

Quines raons pràctiques existeixen per decidir regular la IA a través de la seguretat del producte mitjançant normes harmonitzades? Doncs bé, s'argumenta que els estàndards i processos de certificació s'ajustaran regularment als canvis de la IA i es podran adaptar a les clarificacions respecte a qüestions legals i ètiques en relació amb la IA que encara s'estan debatent (com per exemple l'explicabilitat i la comprensibilitat del sistema)¹³⁹⁶.

Per contra, s'ha argumentat també que els sistemes d'aprenentatge automàtic, poden seguir aprenent mentre funcionen i certes propietats del sistema que s'estableixen en un punt poden no ser vàlides posteriorment. Cal tenir en compte que la natura probabilística de molts sistemes d'IA fa que també sigui difícil establir criteris de qualitat, ja que sovint, estadísticament, no assolixen el 100% de precisió. A més, la qualitat dels sistemes d'IA

¹³⁹⁴ Així i tot, s'ha de fer esment a l'article 47 de la PRIA, que estableix que qualsevol autoritat de vigilància del mercat podrà autoritzar la introducció en el mercat o posada en servei de sistemes específics d'IA de risc alt en l'estat membre del qual es tracti per raons excepcionals de seguretat pública o amb la finalitat de protegir la vida i la salut de les persones, el medi ambient i els actius fonamentals d'indústries i infraestructures. També el considerant 68 de la PRIA, disposa que en determinades ocasions els estats membres podran autoritzar, per motius excepcionals de seguretat pública o en vista a protegir la vida i la salut de les persones físiques i la propietat industrial i mercantil, la introducció en el mercat o la posada en servei de sistemes d'IA que no hagin estat sotmesos a una avaluació de conformitat.

¹³⁹⁵ VEALE, M. i ZUIDERVEEN, F., "Demystifying the Draft EU...", *op. cit.*, 2021, pàg. 102. Igualment l'article 43.1 i 63.4 de la PRIA, estableixen que en el cas de sistemes biomètrics emprats per la policia, autoritats d'immigració o d'asil es requerirà l'avaluació de conformitat d'un òrgan públic.

¹³⁹⁶ EBERS, M., "Standardizing AI. The Case...", *op. cit.*, 2022, pàg. 332. En relació amb adaptar la normativa als canvis en la intel·ligència artificial que resulten del progrés tecnològic cal fer esment a que la Comissió Europea, en diferents preceptes de la PRIA (p. e. art. 4, 7.1, 11.3 o 41.1) es reserva la possibilitat de dictar actes d'execució del Reglament o modificar-ne determinades parts. La regulació d'aquesta delegació s'estableix en els articles 73 i 74 de la PRIA.

mai pot ser avaluada en general, sinó que sempre s'ha d'aplicar el sistema a un conjunt de dades concret, llavors, mentre que els processos d'estandardització estan sempre basats en criteris verificables independentment (per exemple, un full de paper DIN A4 sempre ha de mesurar 210 x 297 mm), els sistemes d'IA manquen d'aquestes mesures de qualitat verificables (per exemple, un sistema d'IA per reconèixer senyals vials amb un grau de certesa molt alt en certes condicions climàtiques -quan fa sol-, pot fallar completament en altres condicions -quan plou-). Addicionalment, com que els sistemes d'IA s'usen en diferents indústries i sectors, amb diferents requisits i característiques, es necessitarà cooperació entre estàndards específics del sector i els estàndards generals. Finalment, els sistemes d'IA són sistemes sociotècnics amb un gran potencial d'impactar en la societat i la qualitat dels quals depèn enormement de la seva aplicació en un àmbit concret, de manera que, perquè no siguin nocius, no serà suficient focalitzar-se en la tecnologia i imposar requeriments únicament a components tècnics, que és el que es pretén fer amb les normes harmonitzades¹³⁹⁷. Un sistema d'IA no pot ser equiparable a un ascensor o a un joc, ja que té implicacions molt més diverses.

Els proveïdors de sistemes d'IA podran, o bé, seguir els estàndards harmonitzats creats per les organitzacions europees de normalització, o bé, interpretar els requisits essencials plantejats en el Reglament de la IA pel seu compte. En el cas de seguir els estàndards harmonitzats hauran de dur a terme un control intern i gaudiran d'una presumpció de conformitat, i en canvi, en el cas de no seguir-los hauran de confiar en les seves pròpies tècniques per assegurar el compliment, havent d'especificar uns requisits vagues assumint els seus propis riscos¹³⁹⁸, i passar, en comptes de l'avaluació de conformitat basada en el control intern, el procediment d'avaluació de la conformitat fonamentat en l'avaluació del sistema de gestió de la qualitat i l'avaluació de la documentació tècnica, amb la participació d'un organisme notificat¹³⁹⁹. Van Gestel i Micklitz estableixen que els estàndards harmonitzats són més barats pels fabricants i una aposta més segura per assolir la comercialització en el mercat del producte¹⁴⁰⁰, per consegüent, Veale i Zuiderveen argumenten que en l'estandardització és on la creació de normes realment ocorrerà¹⁴⁰¹, i

¹³⁹⁷ *Ídem, ibidem.*

¹³⁹⁸ *Ídem*, pàg. 338.

¹³⁹⁹ Article 43.1 de la PRIA.

¹⁴⁰⁰ VAN GESTEL, R. i MICKLITZ, H.W., "European integration through standardization: How judicial review is breaking down the club house of private standardization bodies", *Common market law review*, vol 50, núm. 1, 2013, pàg. 157.

¹⁴⁰¹ VEALE, M. i ZUIDERVEEN, F., "Demystifying the Draft EU...", *op. cit.*, 2021, pàg. 105.

que per tant el CEN i el Cenelec són els actors més importants de la PRIA, malgrat no es mencionin en el text, atès que són les organitzacions europees de normalització a les quals la Comissió pot demanar el desenvolupament d'estàndards harmonitzats¹⁴⁰².

S'alerta llavors que les organitzacions europees de normalització són organismes de dret privat¹⁴⁰³ i que el fet que siguin elles les que acabin determinant els requisits que hauran de seguir els sistemes d'IA per comercialitzar-se, pot suposar una delegació de poders normatius a organismes de dret privat.

4.1.2. Delegació de poders normatius a organismes de dret privat

Els estàndards harmonitzats són normes voluntàries fetes per organismes de dret privat (organitzacions europees de normalització) que es gestionen com organitzacions sense ànim de lucre internacionals en la legislació belga¹⁴⁰⁴. El TJUE mai ha emprat la paraula delegació en les seves sentències¹⁴⁰⁵ mentre que els/les advocats/des generals descriuen el Nou Marc Legislatiu com una delegació normativa, «controlada», a favor d'un organisme de normalització privat¹⁴⁰⁶ o una cessió de fet de les competències a associacions privades¹⁴⁰⁷.

Els estàndards harmonitzats tenen dret preferent en la legislació nacional, és a dir, els estats membres han d'acceptar tots els sistemes d'IA de risc alt que siguin conformes als

¹⁴⁰² *Idem*, pàg. 104.

¹⁴⁰³ *Idem*, pàg. 105; EDRI *et al.*, “An EU Artificial Intelligence Act for Fundamental Rights. A civil Society Statement”, 30 November 2021, pàg. 5. Disponible a: <https://edri.org/wp-content/uploads/2021/12/Political-statement-on-AI-Act.pdf> Darrera visita el 2 de setembre de 2023.

¹⁴⁰⁴ EBERS, M., “Standardizing AI. The Case...”, *op. cit.*, 2022, pàg. 339.

¹⁴⁰⁵ TJUE (Sala Tercera), sentència de 27 d'octubre de 2016, C-613/14, James Elliott Construction Limited v Irish Asphalt Limited, par. 43. En aquest assumpte el TJUE estableix que “*la elaboración de una norma armonizada de ese tipo se confía a un organismo de Derecho privado*”.

¹⁴⁰⁶ Conclusiones del abogado general M. Campos Sanchez-Bordona, presentadas el 28 de enero de 2016, en el asunto C-613/14, James Elliott Construction Limited contra Irish Asphalt Limited, par. 55. En paraules de l'advocat general: “*Tanto la posibilidad de objeción formal por parte de los Estados miembros y del Parlamento Europeo, como la actuación de la Comisión previa a la publicación de las normas técnicas armonizadas ponen de relieve que se trata de un supuesto de delegación normativa, «controlada», a favor de un organismo de normalización privado*”.

¹⁴⁰⁷ Conclusiones de la abogada general Verica Trstenjak, presentadas el 28 de marzo de 2012, en el asunto C-171/11, Fra.bo SpA contra Deutsche Vereinigung des Gas- und Wasserfaches eV (DVGW) – Technisch-Wissenschaftlicher Verein, par. 49. En paraules de l'advocada general: “*elaboración y aplicación de normas técnicas mediante la cesión (de hecho) de las competencias a asociaciones privadas*”.

estàndards harmonitzats¹⁴⁰⁸. La imposició de requisits addicionals en la legislació nacional a productes que estan coberts per estàndards harmonitzats pot suposar que un estat membre infringeixi l'art. 258 TFUE¹⁴⁰⁹. S'ha argumentat, doncs, que els estàndards harmonitzats tenen efectes legals que són propers a les normes legals, podent entendre'ls com una delegació del poder legislatiu¹⁴¹⁰.

El fet que siguin organismes de dret privat els que creïn els estàndards harmonitzats planteja el dubte sobre si el Nou Marc Legislatiu és una delegació il·legal del poder legislatiu de la UE a organismes de dret privat¹⁴¹¹. Sobretot, aquesta qüestió és de vital importància en el cas de la regulació de sistemes d'IA, que sota el nostre parer no poden ser comparables a productes com ascensors o jocs. En els sistemes d'IA hi ha decisions ètiques que no es poden ignorar ni regular mitjançant estàndards harmonitzats¹⁴¹², i la manera en la qual els requeriments legals i els valors ètics es poden traduir en estàndards i especificacions tècniques no està clara actualment¹⁴¹³.

Les organitzacions europees de normalització poden acabar ostentant el poder d'avaluar els interessos en joc -qüestionant-nos fins a quin punt ho poden fer de manera equitativa. Poden dur a terme funcions reguladores, permetre o prohibir certes pràctiques que posen els interessos públics i personals en risc, soscant el poder públic i creant problemes de rendició de comptes. S'ha emfasitzat la manca de supervisió democràtica i participació significativa de les persones afectades i grups d'interès durant el procés de redacció dels estàndards¹⁴¹⁴, tenint en compte que moltes organitzacions civils i associacions de consumidors no tenen cap experiència en estandardització, i inclús poden no estar representades a escala europea¹⁴¹⁵, de manera que no sembla massa realista que els grups d'interès siguin capaços d'influenciar el procés d'estandardització dels sistemes d'IA de

¹⁴⁰⁸ EBERS, M., "Standardizing AI. The Case...", *op. cit.*, 2022, pàg. 339.

¹⁴⁰⁹ TJUE (Sala Desena), sentència de 16 d'octubre de 2014, C-100/13, Comissió Europea contra la República Federal d'Alemanya, par. 63. En aquest assumpte el TJUE estableix que la República Federal Alemanya ha incomplert les obligacions establertes en els articles 4.2 i 6.1 de la Directiva 89/106, al imposar a productes de construcció contemplats per les normes harmonitzades EN 681-2:2000 (...), i que ostenten la marca "CE", exigències suplementaries pel seu accés efectiu al mercat i la seva utilització en el territori alemany.

¹⁴¹⁰ EBERS, M., "Standardizing AI. The Case...", *op. cit.*, 2022, pàg. 339.

¹⁴¹¹ Aquesta cessió sembla respondre a la mitigació del fet que en la regulació pública manqui l'expertesa tècnica.

¹⁴¹² EBERS, M., "Standardizing AI. The Case...", *op. cit.*, 2022, pàg. 339.

¹⁴¹³ *Ídem*, pàg. 331.

¹⁴¹⁴ EBERS, M., "Standardizing AI. The Case...", *op. cit.*, 2022, pàg. 341.

¹⁴¹⁵ VEALE, M. i ZUIDERVEEN, F., "Demystifying the Draft EU...", *op. cit.*, 2021, pàg. 105.

la mateixa manera que ho faria un procés legislatiu ordinari¹⁴¹⁶. Ebers recomana canvis en el marc organitzacional i estructural d'organitzacions de normalització com el CEN o el Cenelec per facilitar un sistema d'estandardització inclusiu que reflecteixi els drets fonamentals i valors de la UE, incloent la protecció de la persona consumidora i garantint a organitzacions civils interessades i afectades drets de participació efectius¹⁴¹⁷. Addicionalment, la supervisió judicial dels estàndards harmonitzats es veu dificultada per obstacles considerables. Malgrat tenir implicacions legals i pràctiques significatives, actualment són immunes a la revisió judicial¹⁴¹⁸.

La delegació de poders associada amb el Nou Marc Legislatiu ja ha estat criticada en el passat per la manca de control democràtic pels òrgans de la UE, la implicació insuficient de grups d'interès i la inexistència de possibilitats de portar els estàndards harmonitzats davant el control judicial. Això suposa problemes particulars en el cas de sistemes d'IA, on les implicacions van més enllà de mers tecnicismes i entren en àrees com la política pública, la salut, la seguretat, els drets fonamentals i la protecció de persones consumidores¹⁴¹⁹. Les organitzacions europees d'estandardització no estan preparades per aquesta tasca, ja que l'estandardització de sistemes d'IA no és únicament una qüestió de decisions tècniques, sinó que hi ha una sèrie de decisions legals i ètiques que s'han de dur a terme socialment i que no poden ser encarregades únicament a organitzacions d'estandardització privades¹⁴²⁰.

4.2. Avaluació de conformitat

L'article 43 de la PRIA estableix que quan el proveïdor hagi utilitzat normes harmonitzades o especificacions comunes¹⁴²¹ per demostrar el compliment dels requisits

¹⁴¹⁶ EBERS, M., "Standardizing AI. The Case...", *op. cit.*, 2022, pàg. 341.

¹⁴¹⁷ *Ídem*, pàg. 343. El Parlament Europeu, proposa com a esmena a la Proposta de Reglament de la IA que presenta la Comissió, afegir en l'article 40.1.c i considerant 61 que s'involucri a totes les parts interessades en el desenvolupament dels estàndards harmonitzats.

¹⁴¹⁸ *Ídem*, pàg. 342.

¹⁴¹⁹ *Ídem*, pàg. 343.

¹⁴²⁰ *Ídem*, pàg. 342.

¹⁴²¹ Les especificacions comunes segons l'article 3.28 de la PRIA són un document, diferent d'una norma, amb solucions tècniques que proposen una forma de complir determinats requisits i obligacions establerts en el Reglament. L'article 41 de la PRIA estableix que la Comissió podrà adoptar-ne mitjançant actes d'execució.

establerts en el Reglament pels sistemes d'IA de risc alt que s'integrin en l'annex III, haurà de realitzar o bé el procediment d'avaluació de conformitat fonamentat en el control intern, o bé el procediment d'avaluació de la conformitat fonamentat en l'avaluació del sistema de gestió de qualitat i l'avaluació de la documentació tècnica, amb la participació d'un organisme notificat¹⁴²².

En el segon apartat de l'article 43, també s'estableix que els sistemes d'IA mencionats en els punts 2 a 8 de l'annex III, es regiran per l'avaluació de conformitat fonamentada en el control intern¹⁴²³. Sembla rellevant que l'apartat sisè del mateix article atorga a la Comissió poders per adoptar actes delegats que modifiquin els apartats primer i segon de l'article per sotmetre els sistemes d'IA de risc alt mencionats en el punt 2 a 8 de l'annex III al procediment d'avaluació de conformitat fonamentat en l'avaluació del sistema de gestió de qualitat i l'avaluació de la documentació tècnica, amb la participació d'un organisme notificat. Podran modificar els apartats tenint en compte l'eficàcia del procediment d'avaluació de conformitat fonamentat en el control intern per prevenir o reduir al mínim els riscos per a la salut, la seguretat i la protecció dels drets fonamentals¹⁴²⁴. Es pot interpretar, llavors, que la Comissió es reserva la possibilitat de sotmetre els sistemes d'IA de risc alt a una avaluació de conformitat més exhaustiva que la que es preveu en primera instància, atès que és molt probable que la que està prevista, al no tenir supervisió, no sigui eficaç per protegir adequadament els drets fonamentals.

La segona tipologia d'avaluació de conformitat es durà a terme quan no s'hagin aplicat normes harmonitzades o especificacions comunes. Els proveïdors podran escollir qualsevol organisme notificat, menys en el cas de sistemes d'IA que prevegin la posada en servei del sistema per autoritats encarregades de l'aplicació de la llei, les autoritats d'immigració i asil, així com les institucions, organismes o agències de la UE, en les quals l'autoritat de vigilància del mercat actuarà com a organisme notificat¹⁴²⁵. Només les aplicacions de risc alt en l'àrea de la identificació biomètrica i la categorització de

¹⁴²² Article 43.1 de la PRIA.

¹⁴²³ Article 43.2 de la PRIA. El Considerant 64 de la norma estableix que el proveïdor es qui, per norma general, ha de dur a terme l'avaluació de la conformitat dels sistemes sota la seva pròpia responsabilitat, amb l'única excepció dels sistemes d'IA que estiguin destinats a emprar-se per la identificació biomètrica remota de les persones. En el cas d'aquests últims, i en la mesura en què no estiguin prohibits, s'ha de preveure que un organisme notificat participi en l'avaluació de la conformitat.

¹⁴²⁴ Article 43.6 de la PRIA.

¹⁴²⁵ Article 43.1 de la PRIA.

persones naturals (punt 1 de l'annex III) requereixen la implicació d'organismes notificats¹⁴²⁶.

Els organismes notificats són organitzacions tècniques independents escollides pel proveïdor, que són típicament firmes de certificació del sector privat, acreditades per l'autoritat notificant¹⁴²⁷ de l'estat membre en qüestió. Per exemple els TÜV, grups alemanys i austríacs, multinacionals amb centenars d'empleats que inspeccionen i auditen en un gran nombre de sectors, o cossos més especialitzats com el Liftinstituut holandès que certifica ascensors¹⁴²⁸. En teoria, els cossos notificats s'anoten en una llista accessible en línia i estan subjectes a estàndards organitzacionals¹⁴²⁹, però a la pràctica hi ha poca informació analítica general i transversal sobre aquests organismes en termes de la seva estructura legal, el seu estatus públic o privat, el nombre de persones que hi treballen o els seus salaris¹⁴³⁰. Igualment, s'apunta que malgrat les pàgines de la PRIA dedicades a l'actuació dels organismes notificats, gairebé no hi ha situacions en les quals els seus serveis siguin requerits. Per la majoria de sistemes d'IA de risc alt els proveïdors poden marcar els seus sistemes com conformes emprant el procediment d'autoavaluació fonamentada en el control intern¹⁴³¹.

Els sistemes d'IA de risc alt es sotmetran a un nou procediment d'avaluació de conformitat sempre que es modifiquin de manera substancial¹⁴³². El considerant 66 de la norma especifica que en el cas de sistemes d'IA de risc alt que segueixen aprenent després de la seva introducció al mercat o posada en servei -és a dir, aquells que adapten automàticament el model en què es duen a terme les seves funcions-, és necessari establir

¹⁴²⁶ VEALE, M. i ZUIDERVEEN, F., “Demystifying the Draft EU...”, *op. cit.*, 2021, pàg. 106.

¹⁴²⁷ Segons l'article 3.19 de la PRIA l'autoritat notificant és l'autoritat nacional responsable d'establir i dur a terme els procediments necessaris per a l'avaluació, designació i notificació dels organismes d'avaluació de la conformitat, així com el seu seguiment.

¹⁴²⁸ VEALE, M. i ZUIDERVEEN, F., “Demystifying the Draft EU...”, *op. cit.*, 2021, pàg. 106.

¹⁴²⁹ *Ídem, ibidem*. Els autors especifiquen que els organismes notificats segueixen estàndards organitzacionals, incloent els EN ISO/IEC 17000, i diferents obligacions legals. Es publiquen en el Diari Oficial de la UE i en la base de dades de la Unió Europea en línia NANDO (New Approach Notified and Designated Organisations), disponible a: <https://ec.europa.eu/growth/tools-databases/nando/index.cfm>. Darrera visita el 2 de setembre de 2023.

¹⁴³⁰ GALLAND J. P., “The Difficulties of Regulating Markets and Risks in Europe through Notified Bodies”, *European Journal of Risk Regulation*, vol. 4, núm. 3, 2013, pàg. 368.

¹⁴³¹ VEALE, M. i ZUIDERVEEN, F., “Demystifying the Draft EU...”, *op. cit.*, 2021, pàg. 106. Cal fer esment a que l'article 33.11 de la PRIA estableix que els organismes notificats participaran directament o mitjançant representació en les organitzacions europees de normalització.

¹⁴³² Article 43.4 de la PRIA. Segons l'article 3.23, una modificació substancial és un canvi en el sistema d'IA posterior a la seva introducció en el mercat o posada en servei que afecti al compliment dels requisits establerts pel Reglament o provoqui la modificació de la finalitat prevista per la qual s'ha avaluat el sistema d'IA en qüestió.

normes que indiquin que no han de considerar-se modificacions substancials els canvis en l'algoritme i en el seu funcionament que hagin estat predeterminats pel proveïdor i s'hagin avaluat en el moment de l'avaluació de conformitat¹⁴³³. L'article 15.3 de la PRIA, relatiu a les obligacions en base a la precisió, la robustesa i la seguretat dels sistemes d'IA de risc alt, estableix que els sistemes d'IA de risc alt que continuen aprenent posteriorment a la seva introducció en el mercat o posada en servei, es desenvoluparan de manera que els possibles biaixos en la informació de sortida deguts al seu ús com a dades d'entrada en futures operacions, és a dir, el que s'anomena circuit de retroalimentació, s'esmenin degudament amb les mesures oportunes¹⁴³⁴.

En vista d'aquesta regulació podem concloure que les obligacions orientades al compliment de l'objectiu de prevenir danys a béns i drets fonamentals cristal·litzen en obligacions de diligència deguda en el marc de límits normatius ja existents¹⁴³⁵. De la mateixa manera, per exemple, en el considerant 71 de l'RGPD, ja analitzat, es convida la persona responsable de tractament a aplicar mesures tècniques i organitzatives apropiades per garantir que s'impedeixin, entre altres, efectes discriminatoris, decantant-se, d'aquesta manera, per instaurar un model de decisió algorítmica orientada a limitar les potencialitats discriminatòries¹⁴³⁶. Així i tot, com ja hem esmentat en relació amb l'article 10 (governança de dades) i l'article 17 de la PRIA (sistema de gestió de qualitat), davant els sistemes d'aprenentatge autònom, sembla molt complicat establir remeis preventius.

Com hem establert anteriorment, la PRIA de manera predominant imposa obligacions a proveïdors i no a usuaris dels sistemes d'IA. S'ha apuntat que mentre alguns riscos dels sistemes d'IA troben la seva causa en com estan dissenyats els sistemes, hi ha riscos significatius que troben la seva causa en com s'usa el sistema. Els proveïdors no poden avaluar l'impacte complet del sistema d'IA de risc alt durant una avaluació de conformitat i per tant, els usuaris haurien de tenir, en vista de la seva responsabilitat pels riscos relatius a com el sistema està essent usat, obligacions en referència als drets fonamentals¹⁴³⁷, que podrien incloure dur a terme avaluacions d'impacte en els drets fonamentals abans d'emprar un sistema de risc alt d'AI, designar les categories de persones o grups de

¹⁴³³ Considerant 66 de la PRIA.

¹⁴³⁴ Article 15.3 de la PRIA.

¹⁴³⁵ AÑÓN ROIG, M.J., "Desigualdades algorítmicas...", *op. cit.*, 2022, pàg. 29.

¹⁴³⁶ ZUDDAS, P., "Decisión algorítmica y discriminación", *op. cit.*, 2022, pàg. 9.

¹⁴³⁷ EDRI *et al.*, "An EU Artificial Intelligence Act for Fundamental Rights. A civil Society Statement", 30 November 2021, pàg. 3. Disponible a: <https://edri.org/wp-content/uploads/2021/12/Political-statement-on-AI-Act.pdf> Darrera visita el 2 de setembre de 2023.

persones que previsiblement es vegin impactades pel sistema, o avaluar l'impacte del sistema d'IA en el medi ambient i en l'interès públic. En aquest sentit, el Parlament Europeu en les esmenes al text presentat per la Comissió proposa que l'article 29 del Reglament, que recull les obligacions dels usuaris de sistemes d'IA d'alt risc, inclogui que han de dur a terme una avaluació de l'impacte del sistema en el context d'ús específic¹⁴³⁸.

4.3. Marcatge “CE” de conformitat

L'article 19 de la PRIA estableix que els proveïdors de sistemes d'IA de risc alt s'asseguraran que el seus sistemes se sotmeten al procediment d'avaluació de conformitat oportú abans la introducció del sistema en el mercat, elaboraran una declaració de conformitat i col·locaran el marcatge “CE” de conformitat¹⁴³⁹.

Segons l'article 3.24 de la PRIA, el marcatge “CE” de conformitat és un marcatge amb el qual un proveïdor indica que un sistema d'IA es conforme amb els requisits establerts en el Reglament -i altra legislació de la UE que harmonitzi les condicions per a la comercialització de productes i prevegi la seva col·locació¹⁴⁴⁰.

En el cas que l'avaluació de conformitat es dugui a terme a través del procediment fonamentat en el control intern, el proveïdor redactarà una declaració UE de conformitat per a cada sistema d'IA i la mantindrà a disposició de les autoritats nacionals competents durant un període de 10 anys després de la introducció del sistema d'IA en el mercat o la seva posada en servei¹⁴⁴¹. En redactar aquesta declaració, el proveïdor assumirà la responsabilitat del compliment dels requisits establerts en el Reglament¹⁴⁴². En el cas que l'avaluació de conformitat es dugui a terme a través del procediment fonamentat en l'avaluació del sistema de gestió de qualitat i l'avaluació de la documentació tècnica, amb la participació d'un organisme notificat, aquest últim expedirà un certificat que no

¹⁴³⁸ EUROPEAN PARLIAMENT, *Amendments adopted by the European Parliament on 14 June 2023...*, *op. cit.*, 14 June 2023.

¹⁴³⁹ Article 19.1 de la PRIA.

¹⁴⁴⁰ Article 3.24 de la PRIA.

¹⁴⁴¹ Article 48.1 de la PRIA.

¹⁴⁴² Article 48.4 de la PRIA.

excedirà d'un període de 5 anys -que podrà prorrogar-se per períodes no superiors a 5 anys¹⁴⁴³. Interessa l'article 45, que disposa que els estats membres vetllaran per tal que existeixi un procediment de recurs contra les decisions dels organismes notificats a disposició de les parts que tinguin un interès legítim en les decisions¹⁴⁴⁴.

El marcatge "CE" anirà seguit del número d'identificació de l'organisme notificat responsable dels procediments d'avaluació de conformitat¹⁴⁴⁵.

Anant més enllà, en el cas de la discriminació algorítmica, i donada la complexitat tècnica característica de l'algoritme, abordar les exigències de tutela antidiscriminària enfront la discriminació algorítmica sembla conduir a promoure o imposar el desenvolupament de sistemes menys esbiaixats i més justos, essent necessaris per implementar i generalitzar les solucions tècniques, l'impuls i promoció de les normes jurídiques¹⁴⁴⁶. Per promoure sistemes més justos, des de la normativa es podria exigir que els sistemes d'IA de presa de decisió en contextos socials complissin certs requisits per obtenir el marcatge "CE", com per exemple: l'establiment d'eines per evitar biaixos i discriminació a través de la seva detecció i prevenció (per exemple, ús de les categories protegides per detectar biaixos en el sistema sense que es puguin reutilitzar les dades per altres finalitats); l'establiment d'eines per interrompre els circuits de retroalimentació que han de ser anticipats, detectats i trencats; l'establiment de sistemes de reconeixement dels danys grupals; la realització, de manera prèvia a la comercialització del sistema d'IA, d'una avaluació d'impacte de discriminació (Discrimination Impact Assessment); l'establiment d'eines per evitar la perpetuació de l'etiquetatge perniciosos de persones; l'especificació de la mètrica i mesures d'equitat que s'emprin en el sistema d'IA; la implicació de les persones que es veuran afectades pel sistema a través de drets de transparència i participació¹⁴⁴⁷; la realització de proves amb el sistema per detectar discriminació¹⁴⁴⁸; l'establiment de mesures tècniques per mitigar la discriminació interseccional i la discriminació per associació i per error; o l'obligació de documentació que acrediti un

¹⁴⁴³ Article 44.1 i 2 de la PRIA.

¹⁴⁴⁴ Article 45 de la PRIA.

¹⁴⁴⁵ Article 49.3 de la PRIA.

¹⁴⁴⁶ SÁEZ LARA, C., "El algoritmo como protagonista de la relación laboral...", *op. cit.*, 2020, pàg. 46.

¹⁴⁴⁷ BERENDT, B., "The AI Act Proposal...", *op. cit.*, 2022, pàg. 49 i 50. En el seu article es refereix a principis que s'haurien d'introduir en la regulació antidiscriminació en comptes d'en els estàndards harmonitzats.

¹⁴⁴⁸ Per a exemples de proves ja realitzades amb sistemes algorítmics, vegeu: EUROPEAN UNION AGENCY FOR FUNDAMENTAL RIGHTS (FRA), *Fundamental Rights Report 2018*, Luxembourg, 6 June 2018., pàg. 63-64.

grau mínim (per exemple, un 95%) de precisió de l'algoritme en diferents grups de població. Tanmateix, per tal que es pugui dur a terme aquest control, s'hauria de passar d'un sistema de notificació prèvia a un sistema d'avaluació prèvia.

4.4. Seguiment posterior a la comercialització o posada en servei del sistema d'IA

També serà el proveïdor del sistema d'IA qui establirà i documentarà un sistema de seguiment posterior a la comercialització de forma proporcionada a la naturalesa de les tecnologies de la intel·ligència artificial i les amenaces pel sistema de risc alt¹⁴⁴⁹. El sistema recopilarà, documentarà i analitzarà de manera activa i sistemàtica dades pertinents proporcionades pels usuaris o a través d'altres fonts i permetrà al proveïdor avaluar el compliment dels requisits establerts en el Reglament¹⁴⁵⁰. El sistema de seguiment es basarà en un pla de seguiment posterior a la comercialització que formarà part de la documentació tècnica¹⁴⁵¹. El proveïdor davant qualsevol incident greu o falla en el funcionament del sistema d'IA que constitueixi un incompliment de les obligacions, en virtut del dret de la UE, notificarà a l'autoritat de vigilància del mercat¹⁴⁵², definida en l'article 3.26 de la PRIA com l'autoritat nacional que duu a terme les activitats i adopta les mesures previstes en el Reglamento (UE) 2019/1020 del Parlamento Europeo y del Consejo, 20 de junio de 2019, relativo a la vigilancia del mercado y la conformidad de los productos¹⁴⁵³ (aquest és el Reglament que estableix un marc per als controls dels productes que entren en el mercat de la UE)¹⁴⁵⁴.

¹⁴⁴⁹ Article 61.1 de la PRIA. Segons l'article 3.25 que defineix el seguiment posterior a la comercialització, es tracta de totes les activitats realitzades pels proveïdors dels sistemes d'IA destinades a recopilar i examinar de forma proactiva l'experiència obtinguda dels sistemes d'IA que introdueixen en el mercat o posen en servei, amb l'objecte de detectar la possible necessitat d'aplicar immediatament qualsevol tipus de mesura correctora i preventiva que resulti necessària.

¹⁴⁵⁰ Article 61.2 de la PRIA.

¹⁴⁵¹ Article 61.3 de la PRIA.

¹⁴⁵² L'article 59.2 de la PRIA estableix que cada estat membre designarà una autoritat nacional de supervisió entre les autoritats nacionals competents que actuarà com a organisme notificant i com a autoritat de vigilància del mercat. L'article 56 de la PRIA constitueix un Comitè Europeu de la Intel·ligència artificial que oferirà assessorament i assistència a la Comissió i autoritats nacionals de supervisió per oferir una aplicació coherent del Reglament, entre altres. Podríem dir que serà un equivalent al Comitè europeu de Protecció de dades.

¹⁴⁵³ Reglamento (UE) 2019/1020 del Parlamento Europeo y del Consejo, 20 de junio de 2019, relativo a la vigilancia del mercado y la conformidad de los productos y por el que se modifican la Directiva 2004/42/CE y los Reglamentos (CE) n.º 765/2008 y (UE) n.º 305/2011. DOUE, L 169/1, 25.6.2019.

¹⁴⁵⁴ Article 62 de la PRIA.

L'article 64 de la PRIA preveu que es concedirà a l'autoritat de vigilància del mercat accés a dades i documentació en el context de les seves activitats, així com ple accés als conjunts de dades d'entrenament, validació i prova empleats pel proveïdor¹⁴⁵⁵. També podran obtenir accés al codi font del sistema d'IA per avaluar la conformitat del sistema de risc alt amb els requisits del Reglament¹⁴⁵⁶. Qualsevol documentació creada o conservada en virtut del Reglament, podrà ser sol·licitada i accessible per les autoritats o organismes públics nacionals encarregats de supervisar o fer respectar les obligacions de la UE en matèria de drets fonamentals, quan la documentació sigui necessària per a l'exercici de les seves competències dins els límits de la seva jurisdicció¹⁴⁵⁷. Si la informació sol·licitada no és suficient, es podrà sol·licitar a l'autoritat de vigilància del mercat d'organitzar proves del sistema d'IA de risc alt a través de mitjans tècnics¹⁴⁵⁸.

L'article 68 disposa que si una autoritat de vigilància del mercat constata algun incompliment de la normativa, per exemple, que no s'ha col·locat el marcatge "CE" de conformitat, demanarà al proveïdor corresponent que ho esmeni i si l'incompliment persisteix adoptarà totes les mesures adequades per restringir o prohibir la seva comercialització¹⁴⁵⁹. L'article 71 de la PRIA estableix un règim de sancions (que haurà de ser detallat pels estats membres) per certes infraccions del Reglament, per exemple, la més greu, una multa de 30.000.000 d'euros o fins el 6% del volum del negoci total anual mundial de l'exercici financer anterior en el cas d'incompliment de la prohibició de certes pràctiques de la intel·ligència artificial o l'incompliment dels requisits establerts en l'article 10 (governança de dades)¹⁴⁶⁰. En la PRIA es dona un rol important a autoritats de vigilància del mercat, que són òrgans públics amb poder per obtenir informació, aplicar multes administratives, retirar productes i obligar intermediaris a cessar l'oferta de productes i/o cooperar per mitigar riscos. La seva principal font d'informació és una cadena de notificació d'obligacions¹⁴⁶¹. Els usuaris de sistemes d'IA hauran de monitoritzar i informar a proveïdors del mal funcionament d'un producte¹⁴⁶² (art. 29.4), i els proveïdors

¹⁴⁵⁵ Article 64.1 de la PRIA.

¹⁴⁵⁶ Article 64.2 de la PRIA.

¹⁴⁵⁷ Article 64.3 de la PRIA.

¹⁴⁵⁸ Article 64.5 de la PRIA.

¹⁴⁵⁹ Article 68 de la PRIA.

¹⁴⁶⁰ Article 71 de la PRIA.

¹⁴⁶¹ VEALE, M. i ZUIDERVEEN, F., "Demystifying the Draft EU...", *op. cit.*, 2021, pàg. 111.

¹⁴⁶² Article 29.4 de la PRIA.

hauran d'informar a les autoritats de vigilància del mercat si el producte després de posar-lo en el mercat revela riscos o no pel compliment del Reglament¹⁴⁶³.

L'article 59.4 de la PRIA estableix que el personal que treballi en les autoritats de vigilància de mercat tindrà suficients competències i coneixements tècnics que inclouran un coneixement profund dels riscos pels drets fonamentals¹⁴⁶⁴. Com hem esmentat anteriorment, les autoritats de vigilància de mercat són autoritats nacionals que s'encarreguen del control de productes que entren en el mercat de la UE (però no de manera específica de productes d'IA). Sembla que la PRIA els dona una càrrega excessiva i inabastable en no ser una autoritat suficientment especialitzada com per fer front al repte de supervisar els productes que impliquin sistemes d'IA que es comercialitzin en el mercat intern de la UE.

A tall de recopilació, els sistemes d'IA comporten qüestions ètiques difícilment traduïbles a estàndards harmonitzats. Atès que les implicacions dels sistemes d'IA van més enllà de mers tecnicismes, els estàndards haurien de ser discutits democràticament i no creats per les organitzacions europees de normalització. L'avaluació sobre si el sistema d'IA assoleix els estàndards la podrà fer el propi proveïdor del sistema en la majoria de casos, tenint la capacitat d'atorgar al seu producte el marcatge "CE" de conformitat, i duent a terme el seguiment posterior a la comercialització. Això suposa que la figura del proveïdor del sistema d'IA ostenta molt poder. Addicionalment, les autoritats de vigilància de mercat actualment són autoritats no especialitzades en IA, fet que suposa que no estigui assegurat un control suficientment exhaustiu de les implicacions dels sistemes d'IA pels drets fonamentals. A més, els usuaris dels sistemes d'IA haurien de tenir, en vista de la seva responsabilitat pels riscos relatius a com el sistema està essent usat, obligacions en referència als drets fonamentals.

5. Conclusions Parcial

¹⁴⁶³ Article 61.1 i 61.2 de la PRIA.

¹⁴⁶⁴ Article 59.4 de la PRIA. El Parlament Europeu en les seves esmenes al text proposat per la Comissió afegeix a l'article l'apartat a) en què s'especifica que cada autoritat nacional de supervisió exercirà els seus poders i les seves tasques de manera independent, imparcial i sense biaixos.

En aquest capítol ens proposàvem respondre la quarta qüestió de recerca del treball, concretament, si el marc normatiu que planteja la UE per regular la intel·ligència artificial ofereix noves garanties per fer front a la discriminació algorítmica. Per respondre a aquesta qüestió hem estudiat la PRIA des de la perspectiva de la regulació de la discriminació algorítmica.

En primer lloc, l'enfocament que planteja la PRIA per regular els sistemes d'IA a través d'una perspectiva de seguretat en el producte col·lideix amb la seva pretesa vocació de protecció genèrica dels drets fonamentals. No s'atorguen drets a les persones interessades (malgrat que això pot canviar en el text final arran de les esmenes proposades per part del Parlament Europeu¹⁴⁶⁵) i la classificació del risc de cada sistema s'estableix d'avançada, fet problemàtic atès que en la discriminació algorítmica el nivell de risc depèn també del context en el qual el sistema es desenvolupa. No obstant això, l'habilitació del tractament de categories especials de dades per garantir la vigilància, detecció i correcció dels biaixos associats al sistema d'IA de risc alt sembla una bona proposta, sempre que s'estableixin salvaguardes reals i efectives i no s'acabi desprotegint els col·lectius més vulnerables.

En segon lloc, si bé és cert que és positiu en relació amb la discriminació algorítmica que es prohibeixi l'ús de certs sistemes d'IA que porten a un tracte perjudicial o desfavorable de certes persones o grups, es fa en un context relativament restringit. Quant als sistemes de risc alt, reconèixer els efectes discriminatoris i danys potencials que poden patir grups oprimits arran de l'ús d'aquests sistemes i no prohibir-los, devalua la importància que la normativa dona a la discriminació. L'establiment de certs requisits, que en la majoria dels casos són autoavaluables pel mateix proveïdor que vol comercialitzar els sistemes, no sembla una protecció suficient. A més, molts sistemes d'IA podran considerar-se de risc baix, i no generaran obligacions, quan potser, aplicats a un context determinat, poden tenir una incidència molt més negativa d'aquella esperada.

En tercer lloc, les exigències de qualitat de les dades que es pretén instaurar és ambiciosa però s'ha criticat que potser és impossible de complir (p. e. bases de dades lliures d'errors i completes). A més, el discurs que es transmet està molt centrat en l'exhaustivitat de la correcció d'errors tècnics per abordar la discriminació, quan l'equitat en els sistemes d'IA depèn també d'altres factors. L'exigència de documentació tècnica, requereix informació

¹⁴⁶⁵ EUROPEAN PARLIAMENT, *Amendments adopted by the European Parliament on 14 June 2023...*, *op. cit.*, 14 June 2023.

de pes relativa als efectes discriminatoris del sistema algorítmic i això suposa que pugui ser potencialment útil en plets de discriminació. Tanmateix, l'accés a aquesta informació està restringit a certes autoritats i no a les persones afectades, limitant la seva incidència. A més, tot i que la supervisió humana s'estableix com un mecanisme repetitiu, continu, que es durà a terme durant tot el cicle de vida del sistema, generarà registres i es desenvoluparà de manera efectiva, per l'experiència prèvia, segueix causant suspicàcies en relació amb la possibilitat que es creï una falsa sensació d'empara sense proveir realment de protecció legal. Quant a la creació i manteniment d'una base de dades de la UE dels sistemes de risc alt descrits en l'annex III, sembla un bona proposta que pot tenir un impacte positiu en la detecció i desafiament de la discriminació algorítmica, malgrat que hauria de ser obligatori el registre de tots els sistemes d'IA i no només aquells de risc alt. Els espais de prova, malgrat que estiguin creats per potenciar la innovació, donen una oportunitat per identificar les vulnerabilitats dels sistema i obren possibilitats en la lluita antidiscriminació, tenint en compte que els organismes d'igualtat nacionals poden estar implicats en el seu funcionament.

Finalment, en quart lloc, els sistemes d'IA comporten qüestions ètiques difícilment traduïbles a estàndards harmonitzats. Atès que les implicacions dels sistemes d'IA van més enllà de mers tecnicismes, els estàndards haurien de ser discutits democràticament i no creats per les organitzacions europees de normalització. A més, la majoria de sistemes d'IA de risc alt podran comercialitzar-se havent passat una autoavaluació. La PRIA de manera predominant imposa obligacions a proveïdors i no a usuaris dels sistemes d'IA (que en vista de la seva responsabilitat pels riscos relatius a com el sistema està essent usat haurien de tenir obligacions en referència als drets fonamentals). Tant pel marcatge "CE" de conformitat com pel seguiment posterior a la comercialització, la figura del proveïdor del sistema d'IA ostenta molt poder. A més, les autoritats de vigilància de mercat actualment són autoritats no especialitzades en IA. Això pot implicar que no hi hagi un control suficientment exhaustiu de les implicacions dels sistemes d'IA per als drets fonamentals.

Per tant, concloem que des de la perspectiva de la protecció de les persones envers la discriminació algorítmica, el nou marc regulador de la IA que proposa la Comissió Europea estableix certes salvaguardes addicionals, com per exemple, l'establiment de certes prohibicions d'ús de la intel·ligència artificial potencialment discriminatòries, l'establiment de certs requisits per a sistemes considerats de risc alt (entre ells l'obligació

de generar documentació tècnica accessible a les autoritats competents), o l'establiment d'una base de dades pública. No obstant això, pretén equilibrar allò social amb llibertats i interessos¹⁴⁶⁶ a través d'un objectiu formal de reducció (dins la mesura del possible segons la finalitat perseguida) del risc de discriminació¹⁴⁶⁷, defugint l'exigència de resultats concrets; les anàlisis de risc es focalitzen primordialment en l'impacte advers de la IA en els individus, i, menys intensament en grups o col·lectius específics, sense abordar en profunditat l'impacte advers en la societat en el seu conjunt¹⁴⁶⁸; les provisions de transparència afegeixen poc a la legislació existent; la regulació té un àmbit d'aplicació extraordinàriament ampli i, es restringeixen els intents legítims dels estats membres per gestionar els impactes socials dels sistemes d'IA en nom del lliure comerç¹⁴⁶⁹.

El marc interpretatiu del risc o de la prevenció de danys a béns i drets fonamentals no resol la qüestió sobre la racionalitat decisòria i vincula les obligacions orientades a aquesta pretensió amb una obligació genèrica de diligència deguda en el marc dels límits normatius ja existents relatius a les garanties i protecció dels drets fonamentals. Entre les conductes considerades de risc alt, tenen una posició privilegiada aquelles que poden donar lloc a resultats discriminatoris bàsicament perquè existeixen biaixos en les dades o en la classificació¹⁴⁷⁰. Per tant, aquesta normativa segueix girant al voltant dels principis estructurals de protecció de dades i del model de responsabilitat proactiva¹⁴⁷¹ i de la mateixa manera que en l'RGPD, les mesures de rendició de comptes de la PRIA poden considerar-se permissives, fet que no afavoreix al clima de confiança en la IA¹⁴⁷² perseguit per la Comissió Europea. Influenciar el desenvolupament de la IA a través d'una política genèrica serà potser una de les úniques funcions que podrien quedar disponibles als poders governamentals, tenint en compte el canvi de dinàmica de poder entre aquests i les grans companyies (vegeu apartat 1 del capítol V), que ocasiona que els primers es

¹⁴⁶⁶ COTINO HUESO, L., “Nuevo paradigma en la garantías...”, *op. cit.*, 2022, pàg. 80.

¹⁴⁶⁷ LÜTZ, F., “Gender equality and artificial intelligence in Europe. Addressing direct and indirect impacts of algorithms on gender-based discrimination”, *ERA Forum*, vol. 23, núm. 1, 2022, pàg. 48.

¹⁴⁶⁸ SMUHA, N. A., “Beyond the individual: governing AI's societal harm”, *op. cit.*, 2021, pàg. 4.

¹⁴⁶⁹ VEALE, M. i ZUIDERVEEN, F., “Demystifying the Draft EU...”, *op. cit.*, 2021, pàg. 112; SANCHO LÓPEZ, M., “Estrategias legales para garantizar los derechos fundamentales...”, *op. cit.*, 2019, pàg. 26. Ja advertia l'autora en relació amb l'RGPD, que més enllà de la voluntat protectora de l'RGPD respecte la privacitat de la ciutadania, no es podia amagar l'objectiu, inclús més important, de crear unes bases per aconseguir un mercat digital únic i evitar que es continuessin produint obstacles per al mercat interior de la UE.

¹⁴⁷⁰ AÑÓN ROIG, M.J., “Desigualdades algorítmicas...”, *op. cit.*, 2022, pàg. 48.

¹⁴⁷¹ COTINO HUESO, L., “Nuevo paradigma en la garantías...”, *op. cit.*, 2022, pàg. 87.

¹⁴⁷² VARGAS CELIX, M. V., “Entre la confianza ciudadana y el impulso tecnológico. Retos jurídicos de la Inteligencia Artificial a la luz de la propuesta de Reglamento de la UE”, *IUS ET SCIENTIA*, vol. 8, núm. 1, 2022, pàg. 187.

trobin en desavantatge, i no està clar fins a quin nivell poden plantejar una amenaça per les grans companyies tecnològiques en el cas que no compleixin les obligacions estipulades en el Reglament¹⁴⁷³. Confiar en una regulació intuïtivament atractiva però ineficaç podria conduir a que el problema subjacent persisteixi, però la presència de la regulació generi la percepció que el problema s'ha resolt¹⁴⁷⁴.

En vista de totes aquestes consideracions (per un esquema d'aquests sis primers capítols, vegeu l'annex del treball), tot seguit, estudiem les possibilitats que ofereix la instauració d'un sistema de protecció col·lectiva-institucional que reculli els punts forts de les normatives fins ara estudiades i que obri possibilitats al desafiament de les decisions preses mitjançant sistemes algorítmics que puguin causar discriminació.

¹⁴⁷³ GUIHOT, M. *et al.*, “Nudging Robots: Innovative...”, *op. cit.*, 2017, pàg. 455.

¹⁴⁷⁴ GREEN, B., “The flaws of policies requiring human oversight...”, *op. cit.*, 2022, pàg. 16.

TERCERA PART. Protecció col·lectiva-institucional envers la discriminació algorítmica

Després d'haver analitzat en la primera part del treball els punts febles de la normativa europea antidiscriminació i de protecció de dades en relació amb la discriminació algorítmica i, en la segona part, els punts forts d'aquestes normatives i la PRIA -que també pot aportar elements a la regulació europea de la discriminació algorítmica-, podem concloure dues premisses. Primerament, que la protecció que es brinda en la UE envers la discriminació algorítmica és insuficient i una de les raons és la perspectiva marcadament individual que caracteritza aquestes regulacions. Segonament, i en conseqüència de la primera premissa, la concepció individual en referència a la litigació no permet en la generalitat dels casos obtenir proves de l'existència de discriminació algorítmica. Per tant, s'han de plantejar noves estratègies reguladores efectives en aquest àmbit¹⁴⁷⁵.

Veiem múltiples similituds entre els casos de discriminació algorítmica i els àmbits en els quals s'ha promogut l'acció de representació, essent l'àmbit del consum el més destacable. La similitud més significativa suposa la dificultat d'obtenció de justícia davant les vulneracions d'interessos supraindividuals i/o vulneracions múltiples de drets individuals. Normalment es tracta de perjudicis que, per la seva magnitud i tipologia estan plurilocalitzats i afecten una diversitat de persones amb certes característiques comunes. Mentre que en el passat una infracció en relació amb la discriminació d'una empresa privada habitualment només afectava a una persona, ara mitjançant l'ús dels mateixos sistemes d'IA per part de diferents empreses pot vulnerar els drets de moltes persones ja sigui en un mateix estat o en diferents estats membres¹⁴⁷⁶, fet que fa urgent una regulació eficaç de les accions col·lectives en matèria antidiscriminació¹⁴⁷⁷. Tanmateix, la discriminació causada per sistemes d'IA usats pels poders públics d'un estat membre, es produiran dins els seus límits territorials i només afectaran a persones de l'estat membre

¹⁴⁷⁵ RIVAS VALLEJO, P., "Análisis desde el derecho antidiscriminatorio", *op. cit.*, 2022, pàg. 436.

¹⁴⁷⁶ GARCÍA ÁLVAREZ, L., "La aplicación en materia ambiental de las acciones de representación para la protección de los intereses colectivos de los consumidores. A propósito de la Directiva (UE) 2020/1828", *Revista Ítalo-Española de Derecho Procesal*, vol. 2, 2020, pàg. 61.

¹⁴⁷⁷ FARKAS, L., "Limited enforcement possibilities under European anti-discrimination legislation -a case study of procedural novelties: actio popularis action in Hungary", *Erasmus Law Review*, vol. 3, núm. 3, 2010, pàg. 184. L'autora justifica la necessitat de promoure l'acció col·lectiva en matèria antidiscriminació atesa la natura sistèmica i estructural que presenta la discriminació.

en qüestió. Aquí es presenta una dualitat d'àmbits territorials en els quals una decisió presa mitjançant sistemes algorítmics pot produir la seva afectació. Aquesta dualitat té una incidència en l'articulació de mitjans processals, que segurament s'hauran de seguir articulant a escala nacional.

En l'actualitat, ens trobem amb múltiples persones afectades per la discriminació algorítmica que no litiguen (pel desconeixement de l'existència de discriminació, la dificultat d'obtenir prova de discriminació, el cost excessiu, l'expectativa incerta en els tribunals contra grans corporacions, etc.), quedant la infracció sense sancionar i els perjudicis sense compensar, impeding alhora que es produeixi llavors un efecte dissuasiu davant l'infractor.

Per com funciona, la discriminació algorítmica s'inclouria en un dels àmbits que necessita l'ús d'estructures i mecanismes del constitucionalisme modern per orientar la governança de l'espai digital, amb la finalitat d'assegurar un millor equilibri entre les relacions jurídiques i la protecció dels drets fonamentals. Cantarini Guerra, en relació amb l'obra de Gunther Teubner "*O direito como sistema autopoietico*" (1989), indica que la formulació tradicional de l'enfocament dels drets fonamentals, des d'una perspectiva individualista de l'equilibri entre els drets individuals d'actors privats en l'esfera digital, estaria desfasada, essent necessari abordar la dimensió col·lectiva-institucional dels drets fonamentals¹⁴⁷⁸. Per la seva banda, Cohen emfasitza que per assegurar la rendició de comptes dels sistemes algorítmics de manera que avanci més enllà de les concepcions subjecte-cèntriques, es necessiten noves competències reguladores i noves tècniques en benefici de les persones afectades per aquests sistemes¹⁴⁷⁹.

A més, Bigman, Wilson, Arnestad, Waytz i Gray afirmen que la discriminació algorítmica causa menys indignació moral que la discriminació humana, i exploren en el seu estudi el que anomenen "dèficit d'indignació algorítmica"¹⁴⁸⁰, que és l'asimetria d'indignació quan es compara la creada per la discriminació algorítmica i la creada per la discriminació humana. Consideren que aquest dèficit d'indignació té com a conseqüència que sigui menys probable que les persones vegin a l'organització que empra el sistema algorítmic

¹⁴⁷⁸ CANTARINI GUERRA, P., "Governanza algorítmica, explicación por medio del diseño y justicia del diseño", *Anales de la Cátedra Francisco Suárez*, núm. 57, 2023, pàg. 126.

¹⁴⁷⁹ COHEN J. E., "Turning Privacy Inside Out", *Theoretical Inquiries in Law*, vol. 20, núm. 1, 2019, pàg. 30.

¹⁴⁸⁰ En anglès, "algorithmic outrage deficit".

com a responsable de la discriminació que se li ha ocasionat, essent més fàcil llavors que la discriminació passi desapercebuda o sense oposició¹⁴⁸¹. No només això, sinó que també implica la debilitació potencial de l'acció col·lectiva per adreçar la discriminació sistèmica¹⁴⁸². Suggereixen que a aquest dèficit d'indignació algorítmica pot contribuir el fet que les persones tinguin menys predisposició a atribuir motivacions negatives a un algoritme. Les persones són percebudes com a capaces d'intenció i antipatia i, per tant, quan discriminen, és més probable que tinguin una motivació perjudicial. En canvi, els algoritmes no es perceben com capaços d'intenció i antipatia, de manera que és menys probable que s'atribueixi el seu resultat a una motivació perjudicial, fet que ocasionaria que les persones se sentissin menys indignades quan els algoritmes provoquin discriminació¹⁴⁸³. També, Van Zoonen conclou que les persones no se senten prou informades i competents per negociar o disputar activament les tecnologies de vigilància¹⁴⁸⁴. El mateix pot passar si presumeixen que se les ha discriminat. A més, com exposa Farkas, existeixen tensions entre la natura de la discriminació i els mecanismes processals a escala europea¹⁴⁸⁵. El fet que els mecanismes que es focalitzen en situacions de discriminació perpetuades en contra d'individus fallin tant a assegurar un tractament igual per a les persones en risc de discriminació sistèmica, com per a altres membres de la població que puguin patir danys col·laterals que neixin d'estructures socials que discriminen, convida a pensar que no és raonable esperar que les persones lluitin contra situacions concretes de discriminació quan les seves vides estan conformades per pràctiques discriminatòries contra les quals estan desarmades¹⁴⁸⁶.

Per consegüent, en la tercera part d'aquest treball responem a la cinquena i última qüestió de recerca, relativa a si és necessària tant la promoció de la representació col·lectiva com una institució pública de referència per fer front a la discriminació algorítmica. Ens centrem llavors en les possibilitats de la protecció col·lectiva-institucional envers la discriminació algorítmica, que combina l'establiment de drets de representació més amplis amb la intervenció proactiva d'organismes de supervisió estatals, millor equipats

¹⁴⁸¹ BIGMAN, Y. E., WILSON, D., ARNESTAD, M. N. WAYTZ, A. i GRAY, K., "Algorithmic Discrimination Causes Less Moral...", *op. cit.*, 2023, pàg. 4-6.

¹⁴⁸² *Ídem*, pàg. 1.

¹⁴⁸³ *Ídem*, pàg. 3.

¹⁴⁸⁴ VAN ZOONEN, L., "Performance and Participation in the Panopticon: Instruments for Civic Engagement with Urban Surveillance Technologies", a JACOBS, G., SUOJANEN, I., HORTON, K. i BAYERL, P., (eds.), *International Security Management. Advanced Sciences and Technologies for Security Applications*, Cham, Springer, 2021, pàg. 253.

¹⁴⁸⁵ FARKAS, L., "Collective actions under European anti-discrimination law", *op. cit.*, 2014, pàg. 25.

¹⁴⁸⁶ *Ídem*, pàg. 26.

per detectar la discriminació dinàmica produïda pels algoritmes i amb més mitjans per obtenir informació sobre el sistema algorítmic, que pot escapar a la percepció de la persona víctima de la discriminació¹⁴⁸⁷. En aquest sentit, veiem dos àmbits rellevants a analitzar, en primer lloc, l'articulació del sistema de representació col·lectiva en matèria de discriminació algorítmica, atès que és un mecanisme que pot maximitzar els punts forts i mitigar els punts febles de les normatives estudiades al llarg del treball (capítol VII); i en segon lloc, la promoció de la protecció envers la discriminació algorítmica des d'una vessant institucional, més específicament a través dels òrgans d'igualtat dels estats membres de la UE, que puguin, entre altres funcions, investigar casos de discriminació algorítmica, interrelacionar persones que es trobin discriminades mitjançant sistemes algorítmics, o representar davant dels tribunals a persones o grups de persones discriminades (capítol VIII).

¹⁴⁸⁷ HAKKARAINEN, J., "Naming something collective does not make it so: algorithmic discrimination and access to justice", *Internet Policy Review. Journal on internet regulation*, vol. 10, núm. 4, 2021, pàg. 17 i 18.

Capítol VII. Protecció col·lectiva: l'acció de representació en matèria de discriminació algorítmica

Com hem vist al llarg d'aquest treball, impugnar situacions individuals de discriminació algorítmica no només és molt difícil sinó que també és molt poc eficient atesa la dificultat de prova. Aquesta dificultat de prova ve causada, principalment, per la dificultat d'accés a informació de rellevància sobre com l'algoritme tracta a determinats grups socials. Si es tingués accés a aquesta informació, es podria comparar la situació d'una persona concreta amb el tracte del sistema algorítmic al grup social al qual pertany i evidenciar que el sistema algorítmic és discriminatori. L'única regulació on trobem l'obligació de crear aquesta informació i documentar-la en relació amb els sistemes algorítmics a escala europea és en la PRIA, en el cas que es tracti d'un sistema de risc alt. A més d'obligar als proveïdors de sistemes de risc alt a fer-la, la documentació serà accessible a les autoritats nacionals competents i organismes notificats¹⁴⁸⁸ i, segons l'article 64.3 i el considerant 79 de la PRIA, a les autoritats o els organismes públics nacionals que supervisin l'aplicació del dret de la UE que protegeix els drets fonamentals, inclosos els organismes d'igualtat¹⁴⁸⁹. Tanmateix, en la PRIA no s'estableixen drets ni individuals ni col·lectius per denunciar la discriminació patida arran d'un sistema algorítmic.

La possibilitat d'accés a aquesta informació únicament per part d'organismes públics ens ha fet pensar en les seves possibilitats per litigar en benefici de persones afectades mitjançant sistemes algorítmics, és a dir, en les possibilitats que obre l'acció col·lectiva en aquest àmbit. La reparació col·lectiva a la Unió Europea està estretament relacionada amb la recerca sobre l'accés a la justícia¹⁴⁹⁰. L'article 19.1 del TUE disposa que els estats membres establiran les vies de recurs necessàries per garantir la tutela judicial efectiva en els àmbits coberts pel dret de la UE i l'article 47 de la CDFUE reconeix el dret a la tutela judicial efectiva com una expressió del principi europeu més general de protecció judicial efectiva. Mentre que aquest últim article s'origina en base a les tradicions constitucionals nacionals (a Espanya, l'article 24 de la Constitució Espanyola¹⁴⁹¹) i dels articles 6 (dret a

¹⁴⁸⁸ Article 11.1 de la PRIA.

¹⁴⁸⁹ Article 64.3 i considerant 79 de la PRIA.

¹⁴⁹⁰ HAKKARAINEN, J., "Naming something collective...", *op. cit.*, 2021, pàg. 11.

¹⁴⁹¹ Article 24 de la Constitució Espanyola. També és rellevant l'article 53.2 de la mateixa norma que disposa que: "*Cualquier ciudadano podrá recabar la tutela de las libertades y derechos reconocidos en el artículo 14 y la Sección primera del Capítulo segundo ante los Tribunales ordinarios por un procedimiento*

un procés equitatiu) i 13 (dret a un procés efectiu) del CEDH, el caràcter obligatori de la CDFUE ha reforçat el seu estatus constitucional en el dret de la UE. Des del Tractat de Lisboa, la protecció judicial efectiva també la trobem en l'article 19.1 del TUE, que, en crear l'obligació d'establir les vies de recurs necessàries per garantir la tutela judicial efectiva en els àmbits coberts pel dret de la UE, pot veure's com a complementari de l'article 47 de la CDFUE¹⁴⁹².

En aquest capítol responem a la primera part de la cinquena qüestió de recerca, a saber, si és necessària la promoció de la representació col·lectiva per fer front a la discriminació algorítmica per tal de millorar l'accés a la justícia de les persones discriminades. Analitzem, doncs, la necessitat d'articular un sistema efectiu de representació col·lectiva per abordar específicament la discriminació algorítmica (1), com els àmbits anteriorment estudiats en aquest treball, a saber, la normativa antidiscriminació (2), la normativa de protecció de dades (3) i la PRIA (4), poden oferir elements per articular una acció col·lectiva en matèria de discriminació algorítmica, sobretot, situant en una millor posició a la part demandant respecte a la prova. Finalment, acabem el capítol amb un apartat de conclusions parcials (5).

1. Per què és necessària l'articulació d'un sistema efectiu de representació col·lectiva enfront conductes discriminatòries causades mitjançant sistemes algorítmics?

Com a conclusió de la primera i segona part del treball, podem veure que sota la normativa actual, les persones que es veuen discriminades mitjançant sistemes algorítmics no tenen eines suficients per obtenir una tutela judicial efectiva. Un dels camins a seguir per fer front a aquesta problemàtica suposa crear i oferir mecanismes d'accés a la justícia que siguin operatius en el context actual. Per assolir aquesta efectivitat en l'àmbit objecte d'estudi, la nostra proposta passa per promoure la representació col·lectiva, comptant amb un recolzament i promoció institucional (com abordarem en el capítol VIII). López Pacheco afirma que les normes i mitjans institucionals en relació amb l'acció col·lectiva

basado en los principios de preferencia y sumariidad y, en su caso, a través del recurso de amparo ante el Tribunal Constitucional."

¹⁴⁹² BENEDI LAHUERTA, S., "Enforcing EU Equality Law through Collective Redress: Lagging behind?", *Common Market Law Review*, vol. 55, núm. 3, 2018, pàg. 798-799.

tant restringeixen (atès que estableixen unes formes correctes i normativitzades per a la defensa de drets, a més de l'ús estratègic que els governs els donen), com obren oportunitats (atès que ofereixen instruments i recursos institucionals que existeixen per ser utilitzats i apropiats per enfortir les causes col·lectives)¹⁴⁹³. Aquesta tensió marca per complet les possibilitats i usos de l'acció col·lectiva i presenta certa complexitat, ja que existeixen oportunitats jurídiques i polítiques nacionals i internacionals amb diferents tipus de combinació i accés¹⁴⁹⁴. Les oportunitats depenen en tot moment del tipus d'accés, és a dir, com d'obertes o tancades són les instàncies institucionals a les demandes d'acció col·lectiva¹⁴⁹⁵.

En aquest apartat abordem, en tot cas en relació amb la discriminació algorítmica i per resseguir els elements essencials que configuren l'acció col·lectiva, en primer lloc, quines raons existeixen per incentivar la UE a crear normes comunes per tal que els estats membres tinguin un mecanisme d'acció col·lectiva en matèria de discriminació algorítmica (1.1), en segon lloc, quin tipus d'interessos i drets a la igualtat i a la no discriminació es poden protegir mitjançant l'acció col·lectiva (1.2), en tercer lloc, qui ha de tenir legitimació per exercir-los (1.3), en quart lloc, amb quin objectiu es pot dur a terme l'acció col·lectiva antidiscriminatòria, ja sigui per cessar una acció que estigui causant discriminació, ja sigui per obtenir una reparació pel perjudici causat a través d'aquesta acció (1.4), i finalment les diferències entre el model opt-in i el model opt-out (1.5).

1.1. Promoció de l'acció de representació en matèria de discriminació algorítmica en la UE

Estudiem en aquest apartat, en primer lloc, les raons per les quals des de la UE s'ha de plantejar un mecanisme d'acció col·lectiva en matèria de discriminació algorítmica (1.1.1), analitzant seguidament la dificultat de crear mecanismes processals harmonitzats

¹⁴⁹³LÓPEZ PACHECO, J. A., "Movilización y acción colectiva por los derechos humanos en la paradoja de la institucionalización", *Estudios Políticos (Universidad de Antioquia)*, vol. 51, 2017, pàg. 62.

¹⁴⁹⁴ *Ídem*, pàg. 69.

¹⁴⁹⁵ *Ídem*, pàg. 70 i 71. L'autor afirma que si bé existeix reconeixement de drets i mecanismes de protecció nacionals en tots els països, la tipologia d'accés defereix substancialment (vegeu, per exemple, l'accés a l'acció col·lectiva a França i a Espanya).

a tots els estats membres de la UE, raó per la qual els mecanismes col·lectius per fer front a la discriminació algorítmica s'articulen de manera pròpia en cada estat membre (1.1.2).

1.1.1. Raons que justifiquen la necessitat d'un sistema efectiu d'acció de representació en matèria de discriminació algorítmica

En l'ordenament jurídic continental europeu, arran dels principis individualistes del dret romà i la doctrina privatista es parteix de l'interès subjectiu de la persona titular. En matèria de drets fonamentals aquest caràcter essencialment subjectiu dels drets s'ha fet cada cop més difícil de sostenir a causa dels canvis i transformacions tecnològics i socials dels segles XX i XXI¹⁴⁹⁶. Malgrat que els drets relacionats amb el dret de les persones consumidores i el dret del medi ambient han portat a que es desenvolupin conceptes, instruments i tècniques de garantia de drets “difusos”, “supraindividuals”, “col·lectius”, “transpersonals”, “grupals”, etc.¹⁴⁹⁷, els sistemes jurídics encara es troben molt endarrerits en relació amb la realitat social i econòmica i qualsevol intent de configurar drets que no s'identifiquen amb l'estructura i instrumentalització individual i privada troba moltes resistències¹⁴⁹⁸.

Tal com els sistemes jurídics no són individuocèntrics perquè sí, sinó que ho són a causa del desenvolupament i disseny institucional, també els remeis processals es focalitzen en els individus i la reparació individual com a forma d'accedir a la justícia¹⁴⁹⁹. En aquest sentit, i com ja hem esmentat anteriorment, aquesta visió individuocèntrica no parteix de qualsevol persona, sinó que es basa en l'ideal d'un home lliure i racional que deté tots els drets legals¹⁵⁰⁰, visió que per exemple, des de la teoria feminista s'ha vist criticada¹⁵⁰¹.

¹⁴⁹⁶ COTINO HUESO, L., “Nuevo paradigma en la garantías...”, *op. cit.*, 2022, pàg. 81.

¹⁴⁹⁷ *Ídem, ibidem.*

¹⁴⁹⁸ GARCÍA ÁLVAREZ, L., “La aplicación en materia ambiental de las acciones...”, *op. cit.*, 2020, pàg. 78.

¹⁴⁹⁹ HAKKARAINEN, J., “Naming something collective...”, *op. cit.*, 2021, pàg. 19.

¹⁵⁰⁰ *Ídem*, pàg. 4.

¹⁵⁰¹ Vegeu, per exemple: BARTLETT, K., “Feminist legal methods”, *Harvard Law Review*, vol. 103, núm. 4, 1990, pàg. 829-888; MACKINNON, C. A., *Feminism unmodified: Discourses on life and law*, Cambridge, Massachusetts, Harvard University Press, 1987.

També, el fet que el marc tradicional del dret processal es basi fermament en la compensació individual dificulta el desenvolupament de mecanismes col·lectius¹⁵⁰².

En el cas de la discriminació algorítmica, és evident que les salvaguardes processals actuals no són suficients per assegurar l'accés a la justícia i existeixen diversos factors que justifiquen la inclusió i la creació de mecanismes efectius dins l'esfera de l'acció col·lectiva per millorar la situació actual¹⁵⁰³, atès que es tracta d'un mecanisme que pretén donar veu a grups socials els drets dels quals s'han vist lesionats dins una societat determinada¹⁵⁰⁴.

En primer lloc, l'impacte de la discriminació és col·lectiu. Com hem estat defensant al llarg del treball, des d'una perspectiva sociològica i psicològica el perjudici discriminatori emergeix a escala grupal¹⁵⁰⁵. També els danys que ocorren no només són individuals sinó que tenen conseqüències en l'estructura de la societat¹⁵⁰⁶. La discriminació és l'acció final conseqüència de sistemes d'opressió. Adreçar aquests fenòmens experimentats col·lectivament només a través de litigació individual sembla inconsistent i proveeix (a vegades) solucions puntuals, però deixa els patrons discriminadors sense canviar¹⁵⁰⁷. Per lluitar envers els sistemes d'opressió, a part de la sensibilització social es requereix tant que existeixi la possibilitat d'obligar al cessament de les conductes discriminatòries com d'obtenir reparació quan aquestes s'han produït. A més, la impugnació de la discriminació algorítmica en base a situacions individuals és molt poc eficient atesa la dificultat d'obtenir prova. Com hem vist, la normativa estudiada fins ara se centra en atorgar drets individuals a les persones afectades (dret antidiscriminació o l'RGPD -llevat l'article 80-) o directament no donar-ne (PRIA) ocasionant que la regulació pugui ser molt inefectiva. Tanmateix, hi ha altres dificultats com la impossibilitat d'encaminar iniciatives individuals per part de la ciutadania¹⁵⁰⁸, les maneres diverses a través de les quals un grup indeterminat de persones pot patir un dany (inclús col·lectius que no es reconeixen en

¹⁵⁰² HAKKARAINEN, J., "Naming something collective...", *op. cit.*, 2021, pàg. 4.

¹⁵⁰³ FARKAS, L., "Collective actions under European anti-discrimination law", *op. cit.*, 2014, pàg. 26; HAKKARAINEN, J., "Naming something collective...", *op. cit.*, 2021, pàg. 4; BENEDI LAHUERTA, S., "Enforcing EU Equality Law...", *op. cit.*, 2018, pàg. 807.

¹⁵⁰⁴ FARKAS, L., "Collective actions under European anti-discrimination law", *op. cit.*, 2014, pàg. 38.

¹⁵⁰⁵ BENEDI LAHUERTA, S., "Enforcing EU Equality Law...", *op. cit.*, 2018, pàg. 793.

¹⁵⁰⁶ HAKKARAINEN, J., "Naming something collective...", *op. cit.*, 2021, pàg. 2 i 3.

¹⁵⁰⁷ BENEDI LAHUERTA, S., "Enforcing EU Equality Law...", *op. cit.*, 2018, pàg. 795.

¹⁵⁰⁸ PLAZA PENADÉS, J., "Aspectos legales del Big Data y la Inteligencia Artificial", a PEDREÑO MUÑOZ, A., MORENO IZQUIERDO, L., PLAZA PENADÉS, J. i IRANZO CABRERA, M., *Big Data e Inteligencia Artificial: una visión económica y legal de estas tecnologías disruptivas*, Fundació Parc Científic Universitat de València, 2019, pàg. 33.

normes materials¹⁵⁰⁹), el desequilibri de poder entre entitats i organitzacions que empen els sistemes algorítmics i les persones que es veuen afectades pels sistemes algorítmics¹⁵¹⁰, o la manca de consciència sobre que la discriminació algorítmica s'està produint¹⁵¹¹.

En segon lloc, les persones que es puguin veure discriminades a través de sistemes algorítmics potencialment poden trobar-se en diversos estats membres de la UE.

En tercer lloc, les persones individualment tenen molta dificultat en l'accés a prova relativa a si han patit discriminació mitjançant sistemes algorítmics.

En quart lloc, l'articulació de mitjans reactius per fer front a la discriminació algorítmica potser serà un dels últims poders que tindran els estats membres de la UE (si la UE comença a fer regulacions de drets màxims en les quals els estats no poden imposar normatives més restrictives, vegeu PRIA), de manera que aquests s'han de potenciar. En aquest sentit, tot seguit volem analitzar com l'articulació de mecanismes col·lectius enfront la discriminació algorítmica s'haurà d'articular a escala nacional atès que no sembla possible una harmonització dels mecanismes dels estats membres impulsada per la UE.

1.1.2. Dificultats d'un mecanisme d'acció de representació en matèria de discriminació algorítmica harmonitzat en els estats membres de la UE

En relació amb la normativa antidiscriminació europea, la contradicció fonamental que trobem és el fet que la discriminació dirigida cap a certs grups de població està prohibida, però en l'àmbit processal, els estats membres no estan obligats a proveir de mecanismes si una víctima individual de discriminació no pot ser identificada¹⁵¹². De manera que no es requereix als estats membres que desenvolupin mecanismes d'acció col·lectiva per individus els drets dels quals (conferits per la normativa antidiscriminació) han estat

¹⁵⁰⁹ HAKKARAINEN, J., "Naming something collective...", *op. cit.*, 2021, pàg. 8.

¹⁵¹⁰ RICHARDS, N. M. i KING, J. H., "Three Paradoxes of Big Data", *Stanford Law Review Online*, vol. 66, 2013, pàg. 45.

¹⁵¹¹ FARKAS, L., "Collective actions under European anti-discrimination law", *op. cit.*, 2014, pàg. 26.

¹⁵¹² BENEDI LAHUERTA, S., "Enforcing EU Equality Law...", *op. cit.*, 2018, pàg. 786.

només vulnerats a escala grupal¹⁵¹³.

A diferència de la normativa europea en matèria de consum¹⁵¹⁴, a través de la Directiva 2020/1828 -que utilitzarem de referència al llarg del capítol- i en matèria de protecció de dades (on es regula la representació col·lectiva en l'article 80 de l'RGPD i s'inclou com a àmbit d'aplicació de la Directiva 2020/1828, però únicament si existeix una relació de consum), la legislació antidiscriminació europea no ha introduït requeriments per desenvolupar la reparació col·lectiva a escala estatal¹⁵¹⁵ (s'indica en les directives antidiscriminació que associacions, organitzacions i persones jurídiques podran iniciar -sota certes condicions- processos per vetllar pel compliment de la Directiva¹⁵¹⁶, però no desenvolupen com).

En la Resolució de Parlament Europeu, de 2 de febrer de 2012, anomenada “*Hacia un planteamiento europeo coherente del recurso colectivo*”, s'exposa la necessitat d'un instrument jurídic horitzontal que inclogui un conjunt comú de principis i ofereixi un accés uniforme a la justícia de la UE mitjançant recursos col·lectius¹⁵¹⁷, com també ho fa la Recomendación de la Comisión, de 11 de junio de 2013, sobre los principios comunes aplicables a los mecanismos de recurso colectivo de cesación o de indemnización en los

¹⁵¹³ *Ídem*, pàg. 807.

¹⁵¹⁴ Volem especificar aquí altres diferències que veiem entre les infraccions de consum i de discriminació algorítmica. En primer lloc, una diferència essencial entre l'afectació dels drets de les persones consumidores i els drets a la no discriminació quan es veuen implicats sistemes algorítmics (com també a la protecció de dades), és que les persones consumidores que han firmat un contracte amb clàusules abusives o amb béns que no estan en conformitat amb el contracte de venda potser no saben els seus drets però, si més no, saben que estan adversament afectades. En segon lloc, en el dret del consum, la defensa col·lectiva de drets o interessos (pluri)individuals homogenis amb la finalitat d'obtenir una indemnització, normalment implica sumes de diners modestes que no són rentables cost-benefici per invertir diners o temps en una reclamació legal individual i per tant, s'uneixen en una única part processal per tal que l'acció sigui més rentable, en canvi, en el cas de la discriminació algorítmica, la unió en una mateixa acció col·lectiva quan s'han vist afectats drets o interessos (pluri)individuals homogenis vindria de la millora en l'establiment d'indicis de discriminació que podria proporcionar. En tercer lloc, els interessos pròpiament grupals en el dret de les persones consumidores pot incloure grups heterogenis de persones, els quals l'única vinculació que tenen és la relació de consum que han tingut amb una empresa determinada. Per contra, això no passarà en una acció col·lectiva d'interessos pròpiament grupals en el dret antidiscriminatori, que sempre estaran formats en base al fet de compartir una (o diverses) característiques protegides per la legislació. Finalment, en quart lloc, el dret de les persones consumidores implicarà sempre a dues parts privades, és a dir, per una banda persones consumidores i per altra banda empreses, a diferència de l'àmbit antidiscriminadori, on la persona o persones discriminades sempre seran particulars i en canvi, podran haver patit discriminació tant per un algoritme emprat per una empresa privada com per l'administració pública (si bé és cert que molt probablement qui hagi dissenyat, desenvolupat i posat en funcionament el sistema algorítmic serà una empresa privada).

¹⁵¹⁵ BENEDI LAHUERTA, S., “Enforcing EU Equality Law...”, *op. cit.*, 2018, pàg. 794.

¹⁵¹⁶ Article 7.2 de la Directiva 2000/43/CE; article 9.2 de la Directiva 2000/78/CE; article 8.3 de la Directiva 2004/113/CE; i article 17.2 de la Directiva 2006/54/CE.

¹⁵¹⁷ PARLAMENTO EUROPEO, *Resolución del Parlamento Europeo sobre «Hacia un planteamiento europeo coherente del recurso colectivo»*, 2011/2089(INI), Bruselas, 2 de febrero de 2012, par. 15.

Estados miembros en caso de violación de los derechos reconocidos por el Derecho de la Unión, que reconeix que la reparació col·lectiva s'hauria d'aplicar horitzontalment i de manera igual tant en el dret de les persones consumidores com en qualsevol altra àrea del dret de la UE on pugui ser útil¹⁵¹⁸.

No és el que s'ha fet amb la Directiva 2020/1828, que contempla únicament l'acció de representació dels interessos col·lectius de les persones consumidores i en certs àmbits determinats que es descriuen en l'annex I, que en relació amb la discriminació només integra el Reglamento (UE) 2018/302 del Parlamento Europeo y del Consejo, de 28 de febrero de 2018, sobre medidas destinadas a impedir el bloqueo geográfico injustificado y otras formas de discriminación por razón de la nacionalidad, del lugar de residencia o del lugar de establecimiento de los clientes en el mercado interior (d'ara en endavant, Reglament 2018/302)¹⁵¹⁹, que prohibeix la discriminació tant directa com indirecta per raó de nacionalitat, lloc de residència o lloc d'establiment en l'accés a béns i serveis¹⁵²⁰. Tanmateix, l'annex I podrà modificar-se per tal d'incloure nous actes de la UE¹⁵²¹, constituint una llista oberta i, els estats membres tenen la competència, com es recull expressament en el considerant 18 de la norma, d'aplicar les disposicions de la Directiva a aspectes diferents d'aquells contemplats en el seu àmbit d'aplicació¹⁵²². No sembla ser això el que farà Espanya, que en el seu Anteproyecto de Ley de acciones de representación para la protección de los intereses colectivos de los consumidores¹⁵²³, no va més enllà del que requereix la norma europea.

¹⁵¹⁸ Considerando 7 de la Recomendación de la Comisión, de 11 de junio de 2013, sobre los principios comunes aplicables a los mecanismos de recurso colectivo de cesación o de indemnización en los Estados miembros en caso de violación de los derechos reconocidos por el Derecho de la Unión. DOUE, L 201/60, 26.7.2013.

¹⁵¹⁹ Reglamento (UE) 2018/302 del Parlamento Europeo y del Consejo, de 28 de febrero de 2018, sobre medidas destinadas a impedir el bloqueo geográfico injustificado y otras formas de discriminación por razón de la nacionalidad, del lugar de residencia o del lugar de establecimiento de los clientes en el mercado interior y por el que se modifican los Reglamentos (CE) n.o 2006/2004 y (UE) 2017/2394 y la Directiva 2009/22/CE. DOUE, L 60/1, 2.3.2018.

¹⁵²⁰ JABŁONOWSKA, A., KUZIEWSKI, M., NOWAK, A. M., MICKLITZ, H-W., PAŁKA, P. i SARTOR, G., *Consumer law and artificial intelligence. Challenges to the EU consumer law and policy stemming from the business' use of artificial intelligence*, European University Institute (EUI), San Domenico di Fiesole, 2018, pàg. 17.

¹⁵²¹ Considerant 17 de la Directiva 2020/1828.

¹⁵²² Considerant 18 de la Directiva 2020/1828.

¹⁵²³ MINISTERIO DE JUSTICIA DE ESPAÑA i MINISTERIO DE CONSUMO DE ESPAÑA, "Anteproyecto de Ley de acciones de representación para la protección de los intereses colectivos de los consumidores", 2023. Disponible a: <https://www.mjusticia.gob.es/es/AreaTematica/ActividadLegislativa/Documents/Anteproyecto%20de%20Ley%20acciones%20representativas.pdf> Darrera visita el 2 de setembre de 2023.

A propòsit d'una harmonització dels mecanismes processals a Europa, la Comissió Europea ha reconegut que, atesa la diversitat dels sistemes nacionals, manca una perspectiva consistent de la reparació col·lectiva a escala europea que pot reduir el gaudi de drets per part de la ciutadania i suposa una aplicació desigual dels drets¹⁵²⁴. Aquesta manca d'aplicació uniforme està lligada a la clàssica desconexió entre drets i remeis (per exemple, els drets estan garantits pel dret de la UE i els processos s'estableixen en la llei estatal)¹⁵²⁵. Delicostopoulos suggereix que garantir l'aplicació efectiva i la primacia dels drets substantius de la UE requereix un grau de primacia processal de la UE que pot coexistir amb la competència processal estatal¹⁵²⁶, tot i així, l'articulació del sistema plantejat per l'autor sembla molt complicada atès que desafia el poder dels estats membres sobre els seus sistemes processals.

A més, mentre que la racionalitat econòmica o de mercat va ser el principal motor darrera l'inici de la normativa europea antidiscriminació, els preàmbuls de les directives antidiscriminació indiquen que van ser concebuts més com a instruments de drets humans que com a eines d'integració de mercat¹⁵²⁷. Existeix una diferència entre el desenvolupament progressiu del dret secundari de mecanismes de reparació col·lectiva en àrees connectades estretament amb el mercat intern i la manca d'aquesta legislació en àrees com la normativa antidiscriminació que tenen una focalització més forta en els drets humans i socials¹⁵²⁸. En aquest sentit, Benedi Lahuerta afirma que en algunes àrees (típicament amb fortes implicacions econòmiques o estratègicament connectades amb el mercat intern -per exemple, protecció de la persona consumidora, protecció de dades o dret de la competència-) el TJUE ha estat més intrusiu en els sistemes processals dels estats membres, definint habitualment l'àmbit de protecció de la normativa europea més àmpliament¹⁵²⁹.

¹⁵²⁴ EUROPEAN COMMISSION, *Public Consultation: Towards a Coherent European Approach to Collective Redress*, SEC(2011)173 final, Brussels, 4 February 2011, par. 10.

¹⁵²⁵ BELJIN, S., "Rights in EU Law", a PRECHAL, S. i VAN ROERMUND, B., (eds.), *The Coherence of EU Law. The Search of Unity in Divergent Concepts*, Oxford, Oxford University Press, 2008, pàg. 91, 109 i 120.

¹⁵²⁶ DELICOSTOPOULOS, J. S., "Towards European Procedural Primacy in National Legal Systems", *European Law Journal*, vol. 9, núm. 5, 2003, pàg. 609.

¹⁵²⁷ BELL, M., "Constitutionalization and EU Employment Law", a MICKLITZ, H-W., (ed.), *Constitutionalization of European private law: XXII/2*, Oxford, Oxford University Press, 2014, pàg. 149 i 150.

¹⁵²⁸ BENEDI LAHUERTA, S., "Enforcing EU Equality Law...", *op. cit.*, 2018, pàg. 812.

¹⁵²⁹ DOUGAN, M., "Who Exactly Benefits from the Treaties? The Murky Interaction Between Union and National Competence over the Capacity to Enforce EU Law", a LEINARTE, E. i ODUDU, O., (eds), *Cambridge Yearbook of European Legal Studies*, vol. 12, 2010, pàg. 90 i 92.

Tenint en compte que la UE estableix drets i certs requisits mínims en relació amb els mecanismes processals que han de garantir els estats membres, sense desenvolupar-los habitualment de manera àmplia, cal remarcar que en matèria antidiscriminació, a Espanya, s'inclou l'acció col·lectiva en l'àmbit de la discriminació -de manera més extensa del que es requereix en la normativa de la UE-, establint l'article 11 bis de la LEC, (modificat recentment per la Llei 15/2022), les normes sobre la legitimació per a la defensa del dret a la igualtat de tracte i la no discriminació.

Volem precisar que ens referim al llarg d'aquest capítol tant a la Directiva 2020/1828 com a l'article 11 bis de la LEC, com també al dret de les persones consumidores, àmbit on està més desenvolupada l'acció col·lectiva. Tot seguit estudiem els interessos que poden ser protegits a través de l'acció col·lectiva.

1.2. Interessos pròpiament grupals (col·lectius o difusos), interessos (pluri)individuals homogenis i l'interès general

En la discriminació algorítmica, el resultat del sistema algorítmic pot perjudicar a una dona en concret, però aquest perjudici ve ocasionat pel tractament massiu de dades que professa un tracte perjudicial cap a les dones com a col·lectiu, no cap a aquesta dona en concret, de manera que és el col·lectiu qui es veu perpètuament afectat. Tenint en compte aquesta consideració, hi ha diversos interessos que es poden protegir mitjançant una acció col·lectiva. Podem distingir els interessos col·lectius, els interessos difusos, els interessos individuals (o (pluri)individuals) homogenis i l'interès general.

Els interessos col·lectius són supraindividuals i indivisibles, i entre les persones titulars d'aquests interessos existeix una vinculació jurídica prèvia a l'ocurrència del dany, és a dir, es tracta d'un grup indeterminat però determinable de persones, ja que els membres de la col·lectivitat poden ser identificables per l'existència d'una relació jurídica prèvia entre ells mateixos o amb l'altra part¹⁵³⁰ (per exemple, situacions on es lesionen els interessos de persones que conformen una associació o un sindicat).

¹⁵³⁰ AGUIRREZABAL GRÜNSTEIN, M. "Algunas precisiones en torno a los intereses supraindividuales (colectivos y difusos)", *Revista Chilena de Derecho*, vol. 33, núm. 1, 2006, pàg. 75.

Els col·lectius són aquells interessos el titular dels quals és un col·lectiu social protegit per les disposicions contingudes en l'ordenament jurídic¹⁵³¹. L'objectiu de la seva tutela jurisdiccional és la protecció d'un interès determinat pertanyent a un col·lectiu, i com a tal, no individualitzable¹⁵³². Per les característiques dels interessos pròpiament grupals, la seva tutela es veurà garantida amb l'acció col·lectiva de cessació, ja que el seu exercici persegueix evitar la continuïtat del dany provocat o impedir l'inici del mateix¹⁵³³.

En els interessos col·lectius, com ja hem dit, el titular, com a col·lectiu concret, preexisteix a la producció de l'afectació concreta, de manera que no s'exigirà una afectació real i similar de totes les persones del col·lectiu com a requisit per a considerar-se membre del grup quant a capacitat i legitimació, sinó que simplement serà suficient el risc potencial d'afectació dels interessos dels membres del col·lectiu a partir d'una determinada conducta¹⁵³⁴. En les accions d'interessos col·lectius es podrà actuar de manera preventiva. El seu objecte és indivisible, ja que tracta d'interessos supraindividuals (és a dir, pertanyen a la col·lectivitat com a grup en si, impeding la fragmentació de l'objecte en tants integrants del grup com existeixin). Això implica, per exemple, que una dona no és titular de l'interès consistent en la protecció de la totalitat de les dones, sinó que únicament és una de les persones integrants en el col·lectiu titular, i com a tal, estarà protegida per l'interès supraindividual. També implica que les persones del grup no poden disposar sobre els interessos, ja que la titularitat l'ostenta el grup com a tal, en tractar-se d'interessos pròpiament grupals. Addicionalment, els interessos de les dones com a grup social no poden quantificar-se econòmicament, fet que sí que succeeix quan es tracta dels interessos individuals d'una d'elles, de manera que l'objecte dels interessos col·lectius no només és indivisible sinó que també inquantificable, impossibilitant una tutela indemnitzadora, atès que l'ordenament jurídic continental es centra en la indemnització real del dany¹⁵³⁵ (qüestió que abordarem en l'apartat 1.4.2 d'aquest capítol).

Els interessos difusos són interessos també supraindividuals i de natura indivisible. En

¹⁵³¹ COROMINAS BACH, S., *La legitimación activa en las acciones col·lectivas*, Tesi doctoral, Universitat de Girona, Girona, 2015, pàg. 32.

¹⁵³² *Ídem*, pàg. 32 i 33.

¹⁵³³ *Ídem*, pàg. 34.

¹⁵³⁴ *Ídem*, pàg. 35.

¹⁵³⁵ *Ídem*, pàg. 40-42.

són titulars persones indeterminades unides per circumstàncies de fet¹⁵³⁶ (per exemple, situacions on es lesiona el medi ambient, on no es pot conèixer exactament la identitat de les persones receptores de la lesió). En el cas de la discriminació algorítmica, estariem davant de drets i interessos difusos quan arran de l'ús d'un sistema algorítmic es veiessin afectades persones pertanyents a un determinat col·lectiu social (per exemple, dones, persones amb discapacitat, persones majors de 60 anys, etc.), que es veurien unides pel fet que s'han vist discriminades per un mateix sistema algorítmic. En canvi, per tal que els drets i interessos per protegir fossin col·lectius i no difusos, les persones que s'han vist discriminades per un sistema algorítmic haurien de compartir una relació jurídica prèvia (haurien d'estar associades o formar part d'alguna entitat que els vinculés jurídicament de forma prèvia, fet que rarament es donarà en la discriminació algorítmica). Les accions per protegir interessos difusos són les que s'han emprat més freqüentment per desafiar la discriminació estructural en la UE i els estats membres¹⁵³⁷. A més, creiem que atesa la dispersió i magnificació dels efectes de la discriminació algorítmica, els interessos difusos tindran especial rellevància en el cas de la discriminació algorítmica.

També es poden defensar a través de l'acció col·lectiva els interessos (pluri)individuals homogenis, que són aquells de natura individual i de titularitat exclusiva però amb un origen fàctic comú¹⁵³⁸ (per exemple, el cas en el qual es produeix un vessament de mercuri que provoca diferents afectacions de salut a les persones que habiten en l'indret del vessament -càncer, malalties dermatològiques-, així com contaminació i mortaldat del bestiar, on cada persona implicada pot tenir una pretensió diferent que és causada per la mateixa circumstància). Són interessos individuals dels membres d'un grup que s'han vist afectats d'una manera quantitativa o qualitativa igual per una mateixa pràctica. Les accions col·lectives busquen permetre a aquestes persones comunament afectades per una pràctica il·lícita, unir-se en una única posició processal i exercir una acció col·lectiva (normalment d'indemnització) per evitar, d'aquesta manera, que la poca entitat de la indemnització en els anomenats “danys massius”¹⁵³⁹ pugui suposar una barrera per a l'accés a la tutela judicial dels afectats. Aquesta col·lectivitat es genera arran d'haver patit un perjudici similar a causa d'una mateixa conducta lesiva dels seus interessos, per tant,

¹⁵³⁶ AGUIRREZABAL GRÜNSTEIN, M. “Algunas precisiones en torno a...”, *op. cit.*, 2006, pàg. 75. Aquesta tipologia d'acció en el món anglosaxó s'anomena “class action”.

¹⁵³⁷ FARKAS, L., “Collective actions under European anti-discrimination law”, *op. cit.*, 2014, pàg. 26.

¹⁵³⁸ AGUIRREZABAL GRÜNSTEIN, M. “Algunas precisiones en torno a...”, *op. cit.*, 2006, pàg. 75.

¹⁵³⁹ Entenem en aquest treball “dany massiu” com un perjudici que afecta a gran quantitat de persones.

no es pot actuar de manera preventiva com en els col·lectius, sinó que només pot comprendre's com a resposta a la producció del dany¹⁵⁴⁰. Cal apuntar que els interessos (pluri)individuals homogenis són interessos individuals i, més enllà de la procedència o no de la tutela col·lectiva quan es veuen afectats per una mateixa conducta, el seu objecte és perfectament divisible i valorable econòmicament, de manera que no impediria que cada persona afectada pogués exercir l'acció individualment¹⁵⁴¹. Normalment, en l'àmbit del consum, aquesta tipologia d'accions de protecció d'interessos (pluri)individuals homogenis es dona quan hi ha un gran nombre d'afectats on l'afectació individual sol caracteritzar-se per ser de molt poca entitat en comparació amb les costes processals i les taxes judicials¹⁵⁴².

En la discriminació algorítmica, per exemple en el cas d'un procés de selecció de feina, totes les persones que s'hagin vist discriminades s'han vist afectades per elements fàctics i causals comuns (haver patit discriminació per part d'un sistema algorítmic), i en el cas de ser-ne conscients, atesa la situació d'asimetria de poder i la dificultat de reunir prova individualment podrien unir-se en una única posició processal i exercir una acció col·lectiva per a la tutela respecte dels perjudicis patits individualment. En aquest cas, la diferència amb els interessos col·lectius o difusos seria que l'acció col·lectiva s'estaria duent a terme amb l'objectiu de defensar diversos interessos individuals que s'han vist afectats de la mateixa manera (per exemple, diverses dones que s'han vist afectades pel mateix sistema algorítmic en el mateix procés de selecció o en el mateix procés d'obtenció d'una ajuda pública), en comptes de defensar els interessos del col·lectiu concret (per exemple, el cessament de l'ús d'un sistema algorítmic perquè és discriminatori envers les dones).

Per tant, en el cas de la discriminació algorítmica, la defensa col·lectiva d'interessos (pluri)individuals homogenis no es justifica per la poca entitat del dany (com en el cas del dret de les persones consumidores i el perjudici econòmic), sinó que es justifica perquè dins l'afectació en base a uns elements fàctics i causals comuns millora la posició probatòria de la persona discriminada en comparació amb la iniciació del plet individualment. És a dir, presentar més d'una persona perjudicada amb una categoria

¹⁵⁴⁰ COROMINAS BACH, S., *La legitimación activa en las acciones col·lectivas*, op. cit., 2015, pàg. 73 i 74.

¹⁵⁴¹ *Ídem*, pàg. 74.

¹⁵⁴² *Ídem*, pàg. 75.

protegida comuna es podria considerar un indicati que el sistema algorítmic és discriminatori, forçant la part demandada a demostrar que les persones que formen part de la demanda col·lectiva van ser perjudicades per altres factors aliens a la categoria protegida que comparteixen. Tanmateix, com hem esmentat anteriorment en el treball, l'èxit de la pretensió en aquests casos podria ser complicat, atès que es podria justificar que l'algoritme no ha considerat positivament a les persones que presenten la demanda a causa d'altres factors no relatius amb la seva pertinença a una categoria protegida. Benedi Lahuerta exposa que en àrees com el dret de la competència o del medi ambient, les demandes col·lectives que tenen èxit solen requerir una anàlisi factual molt complexa que acostuma a ser inaccessible per a les persones individuals, i que habitualment tendeix a crear un equilibri entre risc i recompensa desfavorable per a les persones demandants¹⁵⁴³. Aquesta seria la situació en la discriminació algorítmica.

Finalment, a través d'una acció col·lectiva també es pot defensar l'interès general, que concerneix als valors i objectius que transcendeixen els interessos concrets de persones ciutadanes i grups¹⁵⁴⁴. En el cas de la discriminació algorítmica, cal plantejar-se si més enllà dels interessos pròpiament grupals d'un col·lectiu, no està en joc un interès general, de manera que també sigui possible l'acció popular com a acció per cessar l'activitat il·lícita discriminatòria. Corominas Bach planteja que l'interès general es caracteritza per ser un mer interès en la legalitat, sense precisar cap qualificació subjectiva especial i, en canvi, l'interès pròpiament grupal és un interès personal relatiu a situacions jurídic-subjectives materials tutelades de manera específica¹⁵⁴⁵. Per tant, el mecanisme de tutela dels interessos pròpiament grupals ha de ser una acció col·lectiva de cessació i/o indemnització, atès que es tracta d'accions la situació jurídica de les quals és un interès legítim específic, mentre que l'acció popular es fonamenta en un interès general que no encaixa en les situacions jurídic-subjectives que presenten els interessos pròpiament grupals¹⁵⁴⁶.

Tanmateix, i com veurem que s'ha argumentat en alguns estats membres de la UE, la justificació per dur a terme una acció en base a l'interès general, en el cas de la discriminació algorítmica, implicaria la defensa que la discriminació que es produeix

¹⁵⁴³ BENEDI LAHUERTA, S., "Enforcing EU Equality Law...", *op. cit.*, 2018, pàg. 784.

¹⁵⁴⁴ Interès general. Diccionario panhispánico del español jurídico. 2023. Disponible a: <https://dpej.rae.es/lema/inter%C3%A9s-general>. Darrera visita el 2 de setembre de 2023.

¹⁵⁴⁵ COROMINAS BACH, S., *La legitimación activa en las acciones col·lectivas*, *op. cit.*, 2015, pàg. 38.

¹⁵⁴⁶ *Ídem, ibídem.*

mitjançant sistemes algorítmics reforça i perpetua desigualtats i és contrària a la defensa de valors i objectius socialment justos, més enllà que un col·lectiu social determinat que comparteix una categoria protegida s'hagi vist perjudicat. Cotino Hueso, en aquest sentit, defensa que en l'ús d'IA i l'anàlisi massiva de dades no s'ha de tenir en compte només a l'individu sinó que s'ha de tenir en compte la protecció de la humanitat present i de les futures generacions, és a dir, una perspectiva supraindividual atès que l'afectació a la dignitat o drets concrets no només es limita als usuaris de les noves tecnologies sinó també aquelles persones que no les utilitzen però que són potencialment subjectes de moltes decisions adoptades a través de l'anàlisi massiva de dades i de la IA¹⁵⁴⁷. L'autor, en aquest cas, planteja que el que s'ha de pretendre protegir és l'interès general de manera abstracta. Estudiem, tot seguit, la legitimació activa en l'acció col·lectiva en matèria de discriminació.

1.3. Legitimació activa en l'acció col·lectiva en matèria antidiscriminatòria

La legitimació processal és, dins de cada procés, la consideració especial que atorga la llei a les persones que es troben en una determinada posició respecte l'objecte litigiós i en virtut de la qual s'exigeix, per tal que la pretensió processal pugui ser considerada en relació amb el fons, que siguin aquestes mateixes persones les que figurin com a part¹⁵⁴⁸.

Segons les directives antidiscriminació, la legitimació activa per dur a terme accions col·lectives en aquesta matèria l'ostenten les associacions, organitzacions i altres persones jurídiques que de conformitat amb altres criteris establerts pel dret nacional tinguin interès legítim en vetllar pel compliment de la directiva, en nom de la persona demandant o en el seu suport i sempre amb la seva autorització¹⁵⁴⁹. Un dels problemes

¹⁵⁴⁷ COTINO HUESO, L., "Nuevo paradigma en la garantías...", *op. cit.*, 2022, pàg. 82.

¹⁵⁴⁸ GUASP DELGADO, J. i ARAGONESES ALONSO, P., *Derecho procesal civil*, Pamplona, Civitas, 2005, pàg. 222.

¹⁵⁴⁹ Article 7.2 de la Directiva 2000/43/CE; article 9.2 de la Directiva 2000/78/CE; article 8.3 de la Directiva 2004/113/CE; i article 17.2 de la Directiva 2006/54/CE; CHOPIN, I. i GERMAINE, C., A comparative analysis of..., *op. cit.*, 2022, pàg. 88. En aquest article s'exposa que en alguns estats membres de la UE, la legislació antidiscriminació disposa que les associacions, sindicats i altres organitzacions tenen legitimació en casos específics de discriminació (incloent Àustria, Bèlgica, Bulgària, Croàcia, Xipre, Txèquia, Estònia, França, Alemanya, Grècia, Hongria, Irlanda, Itàlia, Lituània, Malta, Portugal, Romania, Eslovàquia, Eslovènia, Espanya i Suècia), i en canvi, en altres països, no hi ha una provisió específica pels casos de

que s'ha subratllat de l'acció col·lectiva promoguda per associacions privades és el seu finançament i, en el cas que vingui donat per terceres persones, que aquestes no tinguin interessos comercials darrera la promoció de l'acció col·lectiva. Per això, habitualment, la normativa regula quin tipus d'associacions poden iniciar les accions col·lectives i com poden rebre finançament. En matèria d'antidiscriminació, no hi ha una norma europea que reguli aquesta qüestió, de manera que serà a escala estatal on s'establiran els requisits. A Espanya, tanmateix, el finançament de les associacions privades per litigar no està regulat, ni és il·legal¹⁵⁵⁰.

En les directives no es fa cap distinció, però la legitimació activa en les accions de cessació d'una conducta discriminatòria pot ser diferent a la de les accions de reparació.

En l'àmbit de l'acció de cessació, la titularitat dels interessos pròpiament grupals excedeix a la individualitat tradicional, atès que correspon a la totalitat dels membres d'un determinat col·lectiu. Aquesta titularitat no és exclusiva, ja que més enllà de la titularitat del col·lectiu, està la titularitat dels membres que l'integren, atès que són necessaris per a l'existència dels interessos col·lectius (per exemple, una persona que s'ha vist discriminada per la conducta lesiva)¹⁵⁵¹. En ambdós casos es tractarà d'una legitimació ordinària, ja que l'ostentarà qui tingui la titularitat de dret subjectiu material o de la relació jurídica que es discuteix¹⁵⁵². En canvi, en les accions col·lectives reparadores, la legitimació serà ordinària per a cada persona que s'hagi vist perjudicada en l'exercici de la respectiva tutela individual (cada persona que s'hagi vist discriminada per un fet comú), però la legitimació serà extraordinària quan aquesta persona encomani la seva representació a una entitat o organització (atès que aquesta no serà titular d'un dret subjectiu material o de la relació jurídica que es discuteix, sinó que estarà facultada per la llei per exercir drets aliens)¹⁵⁵³. També serà extraordinària si es legitima la persona que

discriminació però hi ha provisions generals que donen legitimació a certes associacions sota certes condicions (incloent Dinamarca, Letònia, Països Baixos i Polònia).

¹⁵⁵⁰ AGULLÓ AGULLÓ, D., "Los contratos de financiación de litigios por terceros (third-party funding) en España", *Revista de Derecho Civil*, vol. IX, núm. 1, 2022, pàg. 185.

¹⁵⁵¹ GUTIÉRREZ DE CABIEDES, P., *La Tutela jurisdiccional de los intereses supraindividuales: colectivos y difusos*, Elcano, Aranzadi, 1999, pàg. 191 i s.; MORENO CATENA, V., *Derecho Procesal Civil. Parte general*, Valencia, Tirant Lo Blanch, 2021, pàg. 99; Tribunal Constitucional (Sala Primera), sentència 47/1990, de 29 de marzo de 1990, FJ 2. El TC disposa que "el interés profesional de promoción y defensa de una categoría de trabajadores, del que puede ser titular no sólo cada uno de ellos individualmente considerados, sino también cualquier asociación o entidad que haya asumido estatutariamente estos mismos fines".

¹⁵⁵² COROMINAS BACH, S., *La legitimación activa en las acciones col·lectivas*, op. cit., 2015, pàg. 134.

¹⁵⁵³ *Ídem*, pàg. 133 i 134.

s'ha vist discriminada per un sistema algorítmic per a l'exercici d'una acció col·lectiva reparadora i s'estableix que l'efecte de cosa jutjada de la sentència recaigui sobre la totalitat del col·lectiu, ja que es fonamentarà en els drets de totes o part de les persones afectades, de manera que s'actuarà en nom propi però en interès de totes les persones afectades¹⁵⁵⁴. Tot seguit estudiem l'acció de cessació i l'acció de reparació.

1.4. Acció de cessació i acció de reparació

En la discriminació algorítmica l'acció de cessació s'exercitaria amb la finalitat de prevenir, fer que es prohibeixi o reiteri una conducta discriminatòria que afecti interessos d'un grup de persones (1.4.1) i, en canvi, l'acció reparadora s'exercitaria per tal que els efectes produïts per una conducta il·lícita desapareguessin, i en el cas que no fos possible, o bé se'n reduïssin les seves conseqüències o bé s'obtingués una indemnització per haver-les patit¹⁵⁵⁵ (1.4.2).

1.4.1. Acció de cessació com a mitjà per a la protecció de les persones discriminades a través de sistemes algorítmics

La finalitat de l'acció de cessació és aconseguir que cessi la conducta discriminatòria, no només en relació amb una persona concreta, sinó en general. Les accions de cessació permeten tant una tutela preventiva com una cessació de la conducta que genera l'afectació. En el cas de la tutela preventiva, quan l'empresa (o administració pública) no hagi dut a terme encara la conducta infractora en contra dels interessos d'un determinat col·lectiu, es demanarà la prohibició o abstenció d'iniciar-la. Aquesta acció s'integra en un marc en el qual s'efectua un control abstracte¹⁵⁵⁶. Així i tot, aquestes accions són excepcionals atès que és extremadament difícil detectar les conductes il·lícites abans que

¹⁵⁵⁴ *Ídem*, pàg. 134. Es refereix al dret de les persones consumidores però que pot ser extrapolable al dret antidiscriminatori.

¹⁵⁵⁵ *Ídem*, pàg. 58.

¹⁵⁵⁶ TJUE (Sala Tercera), sentència de 28 de juliol de 2016, C-191/15, Verein für Konsumenteninformation contra Amazon EU Sàrl, par. 51.

es produeixin, ja que no són evidents fins que afecten una pluralitat dels membres del col·lectiu¹⁵⁵⁷. Nogensmenys, en el cas que una autoritat de supervisió facultada per la PRIA, a propòsit de la documentació tècnica d'un sistema algorítmic que encara no s'ha comercialitzat o de la seva participació en espais de prova (vegeu apartat 2.3.2 del capítol VIII), detecti que el sistema presenta una precisió menys acurada en relació a certs col·lectius determinats de població, podria exercir-la en defensa dels interessos d'aquests col·lectius.

En els casos en els quals encara s'estigui produint l'afectació, l'acció de cessació tindrà un contingut de deixar de fer, mentre que en el cas que ja hagi cessat però existeixi risc de reiteració, el contingut serà de no fer, el que es coneix com a acció de prohibició de reiteració, podent-se garantir el compliment amb el cobrament de multes punitives¹⁵⁵⁸. Per exemple, a Espanya, en els casos de discriminació (a falta de llei més específica que imposi sancions¹⁵⁵⁹), la LEC estableix en l'article 710 la possibilitat de sancionar a qui sigui condemnat¹⁵⁶⁰. De manera que mitjançant l'acció col·lectiva de cessació es podrà demanar tant que es previngui l'ús d'un algoritme que pugui tenir resultats discriminatoris, com que cessi el seu ús, o que es prohibeixi que es torni a emprar. Aquest mecanisme pot ser de molta utilitat atès que d'una acció individual no es podrà pretendre una condemna en cessació que s'estengui a totes aquelles persones presents i futures que es vegin afectades de la mateixa manera¹⁵⁶¹.

En la Directiva 2020/1828 es disposa que l'acció de cessació abasta diferents formes de tutela jurídica destinades a posar fi o prohibir pràctiques lesives, com per exemple, mesures provisionals o definitives per cessar o prohibir una pràctica quan sigui considerada una infracció, que podran incloure que es declari que la pràctica constitueix

¹⁵⁵⁷COROMINAS BACH, S., *La legitimación activa en las acciones col·lectivas*, op. cit., 2015, pàg. 67.

¹⁵⁵⁸ *Ídem*, pàg. 51.

¹⁵⁵⁹ Per exemple, la Ley 15/2007, de 3 de julio, de Defensa de la Competencia, regula l'establiment de sancions als subjectes infractors, el règim del qual es regula en els articles 61-70, conjuntament amb els articles 49-54 de la mateixa norma que regulen el procediment sancionador.

¹⁵⁶⁰ Artículo 710 LEC: “*Condenas de no hacer- 1. Si el condenado a no hacer alguna cosa quebrantare la sentencia, se le requerirá, a instancia del ejecutante por parte del Secretario judicial responsable de la ejecución, para que deshaga lo mal hecho si fuere posible, indemnice los daños y perjuicios causados y, en su caso, se abstenga de reiterar el quebrantamiento, con apercibimiento de incurrir en el delito de desobediencia a la autoridad judicial. Se procederá de esta forma cuantas veces incumpla la condena y para que deshaga lo mal hecho se le intimará por el Secretario judicial con la imposición de multas por cada mes que transcurra sin deshacerlo. 2. Si, atendida la naturaleza de la condena de no hacer, su incumplimiento no fuera susceptible de reiteración y tampoco fuera posible deshacer lo mal hecho, la ejecución procederá para resarcir al ejecutante por los daños y perjuicios que se le hayan causado*”.

¹⁵⁶¹ COROMINAS BACH, S., *La legitimación activa en las acciones col·lectivas*, op. cit., 2015, pàg. 52.

la infracció, o una obligació total o parcial de publicar la resolució sobre la mesura, en la forma en la qual l'òrgan jurisdiccional o l'autoritat administrativa consideri convenient, o una obligació de publicar una declaració de rectificació¹⁵⁶². La possibilitat que l'acció de cessació pugui incloure que es declari que la pràctica constitueix una infracció o s'obligui totalment o parcialment a publicar la resolució sobre la mesura pot ser de summa importància en els casos de discriminació algorítmica. En aquest sentit, interessa com el Tribunal de Bolonya, en la sentència núm. 2949/2020, de 31 de desembre, després de condemnar a Deliveroo al pagament de 50.000 euros per danys en favor de l'organització sindical recurrent, ordena la publicació del text de la sentència en un diari nacional amb la finalitat de garantir la seva màxima visibilitat¹⁵⁶³.

Els efectes d'una sentència relativa als interessos d'un col·lectiu determinat de persones poden beneficiar indirectament, addicionalment a la col·lectivitat en si mateixa, a les persones individualment afectades¹⁵⁶⁴. Tanmateix, la tutela dels interessos col·lectius o difusos mai podrà ser indemnitzadora, ja que el dany que es produeix es materialitza en una afectació d'aquests interessos que són intangibles des del punt de vista econòmic-valoratiu (malgrat això, veurem en el següent apartat que s'ha plantejat la idea del dany moral col·lectiu). En canvi, això no impedeix que els individus que hagin vist els seus interessos afectats pròpiament puguin sol·licitar una tutela indemnitzadora, si bé aquesta última pretindrà exclusivament la tutela d'interessos (pluri)individuals homogenis¹⁵⁶⁵. Per exemple, si una sentència derivada d'una acció col·lectiva de cessació determina que l'algoritme emprat en un procés determinat de selecció d'una empresa és menys precís, i per tant, discriminatori cap a persones de certa ètnia, les persones que hagin participat en aquest procés col·lectiu es podrien beneficiar d'aquesta sentència indirectament.

Tal com hem esmentat, atesa la dispersió i magnificació dels efectes de la discriminació algorítmica, els interessos difusos tindran especial rellevància, i també sembla que l'acció de cessació en tindrà, ja que moltes persones no seran conscients de la discriminació -i per tant no demanaran reparació- i l'objectiu principal (de les associacions i organitzacions que la promoguin) serà aconseguir que l'algoritme discriminatori no pugui

¹⁵⁶² Article 8.1 i 8.2 de la Directiva 2020/1828.

¹⁵⁶³ FERNÁNDEZ SÁNCHEZ, S., "Frank, el algoritmo consciente de Deliveroo. Comentario a la Sentencia del Tribunal de Bolonia 2949/2020, de 31 de diciembre", *Revista de Trabajo y Seguridad Social. CEF*, núm. 457, 2021, pàg. 192 i 193.

¹⁵⁶⁴ COROMINAS BACH, S., *La legitimación activa en las acciones col·lectivas*, op. cit., 2015, pàg. 47.

¹⁵⁶⁵ *Ídem, ibídem*.

seguir operant i es declari que és discriminatori. Igualment, a continuació, estudiem l'acció de reparació.

1.4.2. Acció de reparació com a mitjà per a la protecció de les persones discriminades a través de sistemes algorítmics

L'acció reparadora té la finalitat de compensar als que han patit discriminació i, a més, dissuadir qui l'ha ocasionada de tornar a repetir el seu comportament. En relació amb l'acció reparadora en matèria de discriminació abordem, en primer lloc, en què pot consistir quan es doni la discriminació algorítmica (1.4.2.1), fent en segon lloc certes precisions en relació amb la reparació del dany causat (1.4.2.2).

1.4.2.1. Particularitats de l'acció de reparació en matèria de discriminació algorítmica

L'acció reparadora s'exercita a fi i efecte que, a partir que recaigui la sentència, els efectes produïts per la conducta il·lícita desapareguin o, en cas que no sigui possible, es redueixin les seves conseqüències¹⁵⁶⁶. La pretensió d'una acció reparadora, en el cas dels efectes discriminatoris provats d'un sistema algorítmic, podria ser restaurar la situació d'ent-se a terme de nou el procés de presa de decisió sense tenir-se en compte, en cap cas, la característica protegida, o per obtenir indemnització per cada persona que hagi patit una lesió arran de la presa de decisió discriminatòria. En el cas de la Directiva 2020/1828, com a referent europeu en l'articulació d'accions col·lectives, es parla en tot cas que l'acció reparadora implicarà a persones consumidores individuals afectades per l'acció de representació, que manifestin expressa o tàcitament la seva voluntat de quedar vinculades a l'acció¹⁵⁶⁷ o, quan la mesura restauradora no especifiqui de manera individual les persones consumidores que es puguin beneficiar de les solucions proposades, es delimiti el grup de persones consumidores que puguin beneficiar-se de les

¹⁵⁶⁶ *Ídem*, pàg. 58.

¹⁵⁶⁷ Article 9.2 de la Directiva 2020/1828.

solucions¹⁵⁶⁸.

Un mateix acte pot afectar tant els interessos del col·lectiu com els interessos (pluri)individuals homogenis de les persones que s'han vist subjectes a una decisió mitjançant un sistema algorítmic. Per exemple, si s'ha incomplert una normativa i un algoritme ha causat discriminació, afectant els interessos d'un col·lectiu determinat (per exemple, les dones -interès pròpiament grupal-), posteriorment s'haurà de veure si aquesta conducta ha afectat interessos individuals de mode massiu (per exemple, com ha afectat a cada dona que s'ha vist perjudicada per la conducta -interès (pluri)individual homogeni-). En aquells casos en els quals s'hagin pogut veure afectats interessos (pluri)individuals homogenis i interessos pròpiament grupals, la tutela dels primers dependrà, en la majoria dels casos, de la tutela dels segons¹⁵⁶⁹. La connexió descrita entre ambdós tipologies d'interessos i les respectives accions de tutela, llavors, obliga al tractament conjunt de les respectives legitimacions, raó per la qual les persones que puguin iniciar ambdues accions hauran de ser les mateixes per tal que es pugui procedir a la seva acumulació. En canvi, l'acció col·lectiva reparadora serà completament autònoma respecte a l'acció col·lectiva de cessació quan l'acte no sigui susceptible d'afectar als interessos pròpiament grupals. En el cas de la discriminació algorítmica, creiem que sempre que hi hagi una afectació a interessos (pluri)individuals homogenis hi haurà una afectació a interessos pròpiament grupals, atès que el perjudici que s'ocasioni a una dona, persona amb discapacitat, etc., suposarà que el sistema algorítmic té un biaix envers el col·lectiu afectat, de manera que el perjudici que estarà patint aquesta persona concreta serà la cara visible d'un perjudici que està patint tot el col·lectiu (igual que passa amb la discriminació tradicional, però més evident en el cas de la discriminació algorítmica). Aquí s'ha de tenir en compte, que el grup determinat de població perjudicat pot ser-ho perquè comparteix més d'una categoria protegida, per exemple, un col·lectiu format per dones, d'ètnia gitana, amb discapacitat, o un col·lectiu format per dones, de més de 60 anys, lesbianes. La discriminació interseccional, en les accions col·lectives, és també un repte, ja que s'haurà de trobar la manera de poder detectar quines són les categories protegides combinades que estan causant el perjudici discriminatori (vegeu apartat 2.3.1 del capítol VIII).

¹⁵⁶⁸ Article 9.5 de la Directiva 2020/1828.

¹⁵⁶⁹ COROMINAS BACH, S., *La legitimación activa en las acciones col·lectivas*, op. cit., 2015, pàg. 126.

1.4.2.2. A propòsit del dany moral i la reparació en matèria de discriminació

En el cas de la discriminació, el dany que es produeix principalment és moral¹⁵⁷⁰, tot i que si ens referim, per exemple, a la discriminació patida en un procés de contractació que ha privat una persona d'obtenir una feina amb un salari, podríem parlar de lucre cessant, és a dir, un dany patrimonial. El dany moral que causa el fet de patir discriminació és valorable econòmicament.

En relació amb l'acció reparadora, Coromines Bach afirma que per dur a terme una acció col·lectiva és imprescindible que l'objecte afectat sigui de la mateixa naturalesa, és a dir, que és impossible esgrimir tutela col·lectiva conjunta de danys patrimonials i danys morals, ja que el nexce de causalitat pot ser diferent¹⁵⁷¹. Per tant, en la discriminació algorítmica, si es vol rebre una indemnització individual per la lesió causada per la discriminació, s'haurà de procedir a examinar cas per cas quin ha estat el dany moral produït i només podrà dur-se a terme en el cas de tutela d'interessos o drets (pluri)individuals homogenis -és a dir, quan un conjunt de persones discriminades pel mateix sistema algorítmic decideixin formular la seva acció col·lectivament. Tot i així, González Cazorla planteja la possible existència d'un dany moral col·lectiu (susceptible d'aplicar-se en el cas que es tutelin interessos o drets col·lectius i no difusos, essent necessari que existeixin persones determinades o determinables), independent i autònom d'aquell que pateixen les persones en la seva esfera personal, que afecta a les persones integrants del col·lectiu, lesionant les seves condicions i qualitat de vida¹⁵⁷². Actualment

¹⁵⁷⁰ Vegeu: ALVAREZ VIGARAY, R., "La responsabilidad por daño moral", *Anuario de derecho civil*, vol. 19, núm. 1, 1966, pàg. 81. Malgrat que no hi hagi una definició pacífica de dany moral, aquest autor el defineix per exclusió i per oposició al dany patrimonial com "*aquel perjuicio que no implica una pérdida de dinero, que no entraña para la víctima ninguna consecuencia pecuniaria o disminución de su patrimonio; o también, como aquel daño que no consiste en una pérdida económica o en una falta de ganancia*".

¹⁵⁷¹ COROMINAS BACH, S., *La legitimación activa en las acciones col·lectivas*, op. cit., 2015, pàg. 111.

¹⁵⁷² GONZÁLEZ CAZORLA, F., "Daño moral colectivo en el derecho del consumo chileno: situación actual y proyecciones", *Latin American Legal Studies*, vol. 8, 2021, pàg. 141. En aquest treball l'autor defensa que el dany moral col·lectiu no es tracta de la suma d'afectacions individuals i no poden tenir-se en compte les repercussions particulars patides pels membres de la col·lectivitat, perquè es tracta d'un dany independent i autònom d'aquell que pateixen les persones en la seva esfera personal. Precisa que és necessari arribar a la idea de la dignitat com a afectació d'un bé de la personalitat en un àmbit col·lectiu o, si més no, com la lesió de les condicions o qualitat de vida de les persones, sustentant-se per si mateixa una lesió autònoma que reclama indemnització, sense requerir, per tant, que existeixi patiment de les persones,

aquesta perspectiva està poc estudiada però la plantejarem perquè considerem que en un futur, i atès l'estat actual d'aquesta qüestió, pot guanyar rellevància.

Seguint amb el dany moral, i en el mateix sentit d'abordar una qüestió que en el futur pot guanyar rellevància, creiem pertinent puntualitzar que no obstant que el sistema judicial graviti entorn a la idea d'indemnització econòmica, en moltes ocasions no és la més pertinent, i pot no ser l'única manera ni la més rellevant de reparar el perjudici causat, encara que la quantia permeti compensar els perjudicis soferts per la conducta discriminatòria. En el cas Braathens Regional Aviation AB, el TJUE determina que tota persona que es consideri víctima d'una discriminació té dret a obtenir una tutela judicial efectiva, que no pot entendre's garantida pel mer pagament d'un import pecuniari encara que sigui el que demana la persona demandant, en especial quan l'interès principal d'aquella persona no és econòmic sinó que s'acreditin els fets recriminats al demandat així com la seva qualificació jurídica¹⁵⁷³. També es precisa en la decisió que no es pot considerar que la indemnització compleixi una funció reparadora satisfactòria ni tampoc un efecte realment dissuasiu respecte a l'autor d'una discriminació, incitant-lo a no reiterar el seu comportament, quan s'eviti declarar l'existència de discriminació¹⁵⁷⁴. Maneiro Vazquez exposa que ambdós requisits (reparador i dissuasiu) es complirien en una sentència condemnatòria que vinculés l'abonament de la indemnització a una declaració prèvia de vulneració d'un dret fonamental¹⁵⁷⁵. En aquest sentit, pretenem evidenciar que la indemnització econòmica està molt vinculada a la reparació individual i que, en els casos de discriminació algorítmica, esdevé summament rellevant que es contempli, en qualsevol cas en el que es provi que una empresa o organització ha dut a terme una conducta discriminatòria, la vinculació de l'obligació de cessar l'acció prohibida o de l'abonament de qualsevol indemnització, a la declaració que determinada empresa o organització ha dut a terme una acció discriminatòria i inclús una disculpa pública.

ni que existeixi una lesió específica, sinó que és suficient amb un menysteniment i alteració en les seves condicions de vida causada per la conducta de que ha ocasionat el dany.

¹⁵⁷³ TJUE (Gran Sala), sentència de 15 d'abril de 2021, C-30/19, Diskrimineringsombudsmannen contra Braathens Regional Aviation AB, par. 47.

¹⁵⁷⁴ *Idem*, par. 49. El tribunal estableix que l'obligació de pagar una quantitat de diners no pot garantir un efecte realment dissuasiu respecte l'autor d'una discriminació incitant-lo a no reiterar el seu comportament discriminatori, prevenint així noves discriminacions per la seva part, quan el demandat qüestiona la discriminació però considera més avantatjós, en termes de cost i imatge, abonar la indemnització sol·licitada per la part demandant, evitant així que el jutge nacional declari l'existència de discriminació.

¹⁵⁷⁵ MANEIRO VÁZQUEZ, Y., *La discriminación por asociación...*, op. cit., 2022, pàg. 195.

Tot seguit, per abordar els punts clau de l'acció col·lectiva, cal fer referència al model d'exclusió i al model de participació voluntària en l'acció col·lectiva, que operen independentment del tipus d'acció que s'exerciti.

1.5. Model d'exclusió (opt-in) o model d'inclusió (opt-out)

Amb independència que l'acció sigui de cessació o indemnitzadora, es poden diferenciar dos models en virtut de l'efecte de la cosa jutjada que s'atorgui a la sentència que posi fi al procediment col·lectiu. Si s'estableix que la sentència tindrà una eficàcia de cosa jutjada per a la totalitat dels membres del grup afectat serà un model d'exclusió o opt-out, i en canvi, si l'efecte de cosa jutjada es limita a aquelles persones del grup que hagin manifestat la seva voluntat de formar part del procés col·lectiu es tractarà d'un model d'inclusió, de participació voluntària, opt-in¹⁵⁷⁶. Tant en el model opt-in com en el model opt-out, el que es pretén és que en una situació d'afectació dels drets d'una pluralitat de persones en base a un mateix fet, s'economitzin els recursos i s'evitin resolucions contradictòries.

El model opt-in es planteja com el que proveeix més garanties pels drets de les persones afectades ja que la sentència recaiguda en el procés només afectarà aquelles persones que hagin decidit formar part del grup mitjançant consentiment exprés en la constitució del grup processalment legitimat o mitjançant adhesió posterior, sempre amb caràcter previ a la resolució del procés. És a dir, les persones que s'adhereixen al procés renuncien a formular qualsevol tipus de reclamació posterior, ja sigui a títol individual o col·lectiu. En canvi, les persones que decideixin no adherir-se podran plantejar, més endavant, les seves pretensions tant individualment com col·lectivament, sense que el procés previ les afecti de cap manera¹⁵⁷⁷.

El model opt-out té un abast més altruista i col·lectiu, proper a remeis reguladors i

¹⁵⁷⁶ CHOCHRÓN GIRÁLDEZ, A. M., "El marco jurídico de la Directiva 2020/1828 en relación con la tutela judicial de los intereses colectivos de los consumidores", *Cuadernos de derecho transnacional*, vol. 14, núm. 2, 2022, pàg. 291; COROMINAS BACH, S., *La legitimación activa en las acciones col·lectivas*, op. cit., 2015, pàg. 136.

¹⁵⁷⁷ COROMINAS BACH, S., *La legitimación activa en las acciones col·lectivas*, op. cit., 2015, pàg. 137.

propostes de justícia social per evitar conductes futures no desitjables. Sens dubte, l'avantatge d'aquest sistema és l'efecte dissuasiu producte de la inclusió de la totalitat del col·lectiu afectat per la decisió o resolució que posi fi a un procés col·lectiu, excepte per aquelles persones que hagin manifestat la seva voluntat expressa de reservar la seva pretensió per a una tutela individual posterior. No obliga les persones afectades a participar en el procés en el cas que no vulguin formar-ne i, per tant, no es requerirà participació de les persones afectades per l'eficàcia de l'acció col·lectiva, sinó únicament de l'entitat legitimada, sense impedir la seva intervenció¹⁵⁷⁸.

La qüestió de l'elecció entre el model opt-in o opt-out, en relació amb la formació del grup i l'efecte de les resolucions, ha estat sempre molt polèmica a propòsit de l'harmonització de remeis col·lectius a la UE, per la suposada contrarietat de les accions opt-out amb l'ordre públic processal d'alguns estats membres i el risc associat d'abus processal d'aquest model de litigació col·lectiva. Tanmateix, l'opció opt-out ha provat els seus avantatges, entre altres, la major taxa de participació en evitar el biaix de la inèrcia, i la major eficàcia, ja que l'increment de participació aconsegueix un major èxit de la reparació i un millor efecte dissuasiu de conductes potencialment lesives¹⁵⁷⁹.

En qualsevol cas, és evident que perquè una acció de representació resulti eficaç és necessari comptar amb un bon mecanisme de publicitat i informació¹⁵⁸⁰ que permeti arribar al nombre més alt possible de persones amb interès en el litigi que podrien voler unir-se a l'acció col·lectiva¹⁵⁸¹. Wachter afirma que la notificació a les persones que conformen un grup, les dades de les quals s'usen en l'elaboració de perfils o la presa de decisions automatitzades, podria facilitar la reclamació de discriminació en base a accions col·lectives¹⁵⁸².

Per exemple, un sistema de publicitat i notificació a les persones afectades per una discriminació algorítmica, adaptat als sistemes telemàtics actuals podria constar de bases de dades electròniques d'accés públic en forma de pàgina web, per informar a la

¹⁵⁷⁸ *Ídem*, pàg. 138.

¹⁵⁷⁹ GARCÍA ÁLVAREZ, L., "La aplicación en materia ambiental de las acciones...", *op. cit.*, 2020, pàg. 98.

¹⁵⁸⁰ CHOCHRÓN GIRÁLDEZ, A. M., "El marco jurídico de la Directiva 2020/1828...", *op. cit.*, 2022, pàg. 288.

¹⁵⁸¹ GARCÍA ÁLVAREZ, L., "La aplicación en materia ambiental de las acciones...", *op. cit.*, 2020, pàg. 100.

¹⁵⁸² WACHTER, S., "Affinity profiling and discrimination...", *op. cit.*, 2020, pàg. 419.

ciutadania de les entitats que estan habilitades per presentar accions nacionals i/o transfrontereres, i també, per proporcionar informació general sobre les accions de representació que estan en curs o que han conclòs¹⁵⁸³, amb l'obligació de publicar no únicament la intenció de dur a terme l'acció sinó també l'estat de la tramitació processal i el resultat obtingut¹⁵⁸⁴. Això últim per evitar tant altres accions col·lectives com individuals amb el mateix objecte i causa contra la mateixa part demandada¹⁵⁸⁵. Aquesta publicació es podria fer a la pàgina web de qui duu a terme l'acció col·lectiva, incloent dades, en llenguatge senzill i comprensible, sobre l'objecte, conseqüències legals possibles, descripció del grup, qüestions respecte la salvaguarda de proves, etc.¹⁵⁸⁶

A tall de recapitulació, en aquest apartat hem presentat els elements bàsics de l'acció col·lectiva (interessos, legitimació, objectiu i models de participació) i la seva adequació en casos de discriminació algorítmica. L'acció col·lectiva en matèria de discriminació algorítmica podria dur-se a terme per protegir interessos col·lectius, difusos, (pluri)individuals homogenis o l'interès general i podria tenir l'objectiu tant de cessar l'acció discriminatòria com de reparar els perjudicis causats. Tanmateix, pel funcionament de la discriminació algorítmica, el més factible sembla ser l'acció col·lectiva de cessació en defensa d'interessos difusos. Sense un sistema harmonitzat a la UE en matèria de discriminació, són les normatives dels estats membres les que determinen l'abast de l'acció col·lectiva, de manera que és a escala nacional on es pot determinar l'abast i la configuració d'aquest tipus d'acció.

Per concretar quines poden ser les aportacions de l'acció col·lectiva per a la discriminació algorítmica, analitzem seguidament en cada àmbit d'estudi d'aquest treball com es pot configurar i quines avantatges pot proporcionar per millorar la situació actual, començant per la normativa antidiscriminació.

¹⁵⁸³ Article 14 i Considerant 63 de la Directiva 2020/1828.

¹⁵⁸⁴ Article 13 de la Directiva 2020/1828.

¹⁵⁸⁵ Article 9.4 de la Directiva 2020/1828.

¹⁵⁸⁶ GARCÍA ÁLVAREZ, L., "La aplicación en materia ambiental de las acciones...", *op. cit.*, 2020, pàg. 101.

2. Normativa antidiscriminació i acció de representació en matèria de discriminació algorítmica

En aquest apartat estudiem quines normes tant europees (2.1), com espanyoles (2.2) habiliten la possibilitat d'iniciar una acció col·lectiva quan moltes persones han patit discriminació causada per un mateix fet. Seguidament analitzem els avantatges que brinda el fet de dur a terme una acció col·lectiva en comptes d'una acció individual en el cas de la discriminació algorítmica (2.3), arran de la jurisprudència establerta per part del TJUE en accions col·lectives relatives a la discriminació (no algorítmica).

2.1. Acció de representació a la UE en matèria de discriminació

En aquest apartat analitzem, en primer lloc, com s'articula l'acció col·lectiva en matèria de discriminació en el si de la UE (2.1.1), i en segon lloc, l'ús en certs estats membres de la UE de l'acció popular per defensar interessos de determinats col·lectius de població que s'han vist discriminats (2.1.2).

2.1.1. Normativa antidiscriminació de la UE que habilita l'acció de representació per discriminació algorítmica

L'acció col·lectiva es defineix com qualsevol mecanisme que pot aconseguir la cessació o la prevenció de pràctiques empresarials (o de l'administració pública) il·legals que afecten moltes persones o la compensació pel perjudici causat per aquestes pràctiques¹⁵⁸⁷. El requeriment de proveir mecanismes de cessació o compensació generalment s'interpreta en línia amb el dret de la UE a una tutela judicial efectiva (art. 47 CDFUE) i l'article 6 del CEDH¹⁵⁸⁸. Aquests mecanismes són procediments judicials, gairebé

¹⁵⁸⁷ EUROPEAN COMMISSION, *Public Consultation: Towards...*, *op. cit.*, 4 February 2011, par. 7.

¹⁵⁸⁸ BENEDI LAHUERTA, S., "Enforcing EU Equality Law...", *op. cit.*, 2018, pàg. 787.

judicials i administratius que assoleixen els requeriments generals del dret de la UE i el CEDH per a una tutela efectiva¹⁵⁸⁹.

Cal puntualitzar que, pel que fa al TEDH i la seva acceptació d'accions col·lectives, l'article 34 del CEDH¹⁵⁹⁰ disposa que una persona, Organització No Governamental (d'ara en endavant, ONG) o grup de persones poden presentar una demanda només si poden provar que ha estat víctima d'una violació dels drets estipulats en el Conveni¹⁵⁹¹. De manera que les accions per a la protecció d'interessos difusos o accions populars no són acceptades en el TEDH.

En relació amb el dret de la UE, per una banda, l'article 7.2 de la Directiva 2000/43/CE, 9.2 de la Directiva 2000/78/CE, l'article 8.3 de la Directiva 2004/113/CE i 17.2 de la Directiva 2006/54/CE estableixen que els estats membres vetllaran per tal que les associacions, organitzacions i altres persones jurídiques que, de conformitat amb els criteris establerts en el dret nacional, tinguin un interès legítim en vetllar pel compliment d'allò disposat en la Directiva, puguin iniciar en nom de la persona demandant o en el seu suport i la seva autorització qualsevol procediment judicial o administratiu previst per tal d'exigir el compliment de les obligacions de la Directiva¹⁵⁹². També la Directiva 2014/54/UE del Parlamento Europeu y del Consejo, de 16 de abril de 2014, sobre medidas para facilitar el ejercicio de los derechos conferidos a los trabajadores en el contexto de la libre circulación de los trabajadores, en un article similar al de les Directives antidiscriminació, estableix que els estats membres han de desenvolupar mecanismes de reparació col·lectiva per adreçar la discriminació de les persones treballadores per raó de nacionalitat¹⁵⁹³.

¹⁵⁸⁹ *Ídem, ibídem.*

¹⁵⁹⁰ Article 34 del CEDH disposa que: “*El Tribunal pot conèixer d'una demanda presentada per qualsevol persona física, organització no governamental o grup de particulars que es consideri víctima d'una violació per una Alta Part contractant dels drets que reconeix el Conveni o els seus protocols. Les Altes Parts contractants es comprometen a no obstaculitzar per cap mitjà l'exercici eficaç d'aquest dret*”.

¹⁵⁹¹ FARKAS, L., “Limited enforcement possibilities under European anti-discrimination...”, *op. cit.*, 2010, pàg. 187; ORGANISATION FOR SECURITY AND CO-OPERATION IN EUROPE (OSCE), *Use of actio popularis...*, *op. cit.*, November 2016, pàg. 8.

¹⁵⁹² Article 7.2 de la Directiva 2000/43/CE; article 9.2 de la Directiva 2000/78/CE; article 8.3 de la Directiva 2004/113/CE; i article 17.2 de la Directiva 2006/54/CE. Per una anàlisi detallada de la legitimació de les organitzacions per actuar en benefici o en suport de les víctimes en casos de discriminació en cada estat membre de la UE, vegeu: CHOPIN, I. i GERMAINE, C., *A comparative analysis of...*, *op. cit.*, 2022, pàg. 92-94.

¹⁵⁹³ Article 3.2 de la Directiva 2014/54/UE del Parlamento Europeu y del Consejo, de 16 de abril de 2014, sobre medidas para facilitar el ejercicio de los derechos conferidos a los trabajadores en el contexto de la libre circulación de los trabajadores. DOUE, L 128/8, 30.4.2014.

Els articles, al referir-se a “*iniciar en nom de la persona demandant o en el seu suport i autorització*” exigeixen que les víctimes estiguin identificades per dur a terme una acció grupal per part d’una associació, organització o entitat jurídica. Tanmateix, en el cas Feryn, el TJUE accepta la possibilitat que els estats membres incloguin l’opció que entitats legitimades litiguin sense necessitar l’autorització de la/les persona/es afectada/es, com per exemple, la legislació belga, que reconeix a un organisme concret, el Centrum voor gelijkheid van kansen en voor racismebestrijding, legitimació activa quan existeixi o pugui existir una discriminació sense necessitar una denúncia prèvia¹⁵⁹⁴. Per tant, els estats membres poden oferir més alternatives i crear règims més protectors¹⁵⁹⁵.

En relació amb la indemnització pel perjudici causat, l’article 18 de la Directiva 2006/54/CE i 8.2 de la Directiva 2004/113/CE estableixen que els estats membres introduiran en els seus ordenaments jurídics interns les mesures necessàries per tal que qualsevol persona que es consideri perjudicada a causa d’una discriminació en el sentit de la Directiva, rebi una indemnització o compensació reals i efectives de l’estat membre, de manera dissuasiva i proporcional al dany patit, sense limitar-la en base a la fixació d’un límit màxim predeterminat¹⁵⁹⁶. Aquestes Directives van més enllà que les Directives 2000/43/CE i 2000/78/CE que només preveuen la indemnització a la víctima a través del règim de sancions per l’incompliment de la Directiva, en els articles 15 i 17 respectivament¹⁵⁹⁷.

En l’assumpte Baathens Regional Aviation AB, ja esmentat, el TJUE va precisar certes qüestions sobre la indemnització per discriminació en relació amb la Directiva 2000/43. La indemnització per discriminació correspon a una sanció en el sentit de l’article 15 de la Directiva 2000/43/CE, i ha de determinar-se en cada cas concret, de manera que constitueixi una indemnització raonable per a la víctima i ajudi a la lluita envers les discriminacions en la societat, assumint així la seva doble funció de reparació i prevenció¹⁵⁹⁸. Quan la reparació pecuniària és la mesura adoptada, ha de ser adequada en

¹⁵⁹⁴ TJCE, sentència de 10 de juliol de 2008, C-54/07, Centrum voor gelijkheid van kansen en voor racismebestrijding contra Firma Feryn NV, par. 14.

¹⁵⁹⁵ FARKAS, L., “Limited enforcement possibilities under European anti-discrimination...”, *op. cit.*, 2010, pàg. 184.

¹⁵⁹⁶ Article 18 de la Directiva 2006/54/CE i article 8.2 de la Directiva 2004/113/CE.

¹⁵⁹⁷ Article 15 de la Directiva 2000/43/CE i article 17 Directiva 2000/78/CE. Articles amb un contingut similar els trobem també en la Directiva 2006/54/CE (art. 28) i la Directiva 2004/113/CE (art. 14).

¹⁵⁹⁸ TJUE (Gran Sala), sentència de 15 d’abril de 2021, C-30/19, Diskrimineringsombudsmannen contra Braathens Regional Aviation AB, par. 24.

el sentit que ha de permetre que es compensin els perjudicis íntegrament patits a causa de la discriminació, segons les normes nacionals aplicables. En canvi, una sanció merament simbòlica no pot considerar-se compatible amb l'aplicació correcta i eficaç de la Directiva 2000/43/CE¹⁵⁹⁹.

Per altra banda, com ja hem esmentat, en l'àmbit del dret de les persones consumidores, la Directiva 2020/1828 únicament integra en el seu àmbit d'aplicació el Reglament 2018/302, sense incloure altres formes de discriminació¹⁶⁰⁰. Sota el nostre parer, la inclusió de totes les formes de discriminació que es poden donar en les relacions de consum podria haver millorat la protecció a escala europea de la discriminació algorítmica, ja que hauria obligat a tots els estats membres a desenvolupar i ampliar les eines d'acció col·lectiva en casos de discriminació en l'àmbit del consum.

2.1.2. Acció popular a Europa en defensa dels interessos de col·lectius socials que es veuen discriminats

Cal també fer èmfasi en la diferència de conceptualització entre els interessos a protegir a través de l'acció col·lectiva en els estats membres en no estar establert per part de la UE un mecanisme comú. Per exemple, malgrat que seguint el criteri que defensa Corominas Bach, l'acció popular pretén l'interès general, entenent l'interès general com un mer interès en la legalitat, sense precisar cap qualificació subjectiva especial¹⁶⁰¹, és a dir, sense la capacitat de defensar grups socials específics (vegeu apartat 1.2 d'aquest capítol), l'Organització per la seguretat i cooperació a Europa (OSCE) relata com alguns estats membres de la UE fan servir l'acció popular per possibilitar l'acció col·lectiva de certs organismes públics i associacions en defensa de l'interès d'un col·lectiu determinat de la població (el que en l'ordenament jurídic espanyol seria la defensa d'interessos difusos). En els ordenaments jurídics d'aquests estats membres s'entén l'acció popular com un mecanisme per a la protecció d'un grup particular de persones en contra les violacions

¹⁵⁹⁹ *Ídem*, par. 38.

¹⁶⁰⁰ Entenem que és així ja que en l'article 2.1 de la Directiva 2020/1828 es disposa que s'aplica a les accions de representació exercitades enfront d'actes d'empresaris que infringeixin les disposicions del dret de la UE recollides en l'annex I, i en l'annex I en relació amb la discriminació només es recull el Reglament 2018/302.

¹⁶⁰¹ COROMINAS BACH, S., *La legitimación activa en las acciones col·lectivas*, op. cit., 2015, pàg. 38.

sistèmiques de drets que representen un interès públic en una societat definida com a democràtica¹⁶⁰². Posen l'exemple de països de la UE com Croàcia, Hongria i Eslovàquia¹⁶⁰³. Farkas se centra en el cas d'Hongria, on la normativa antidiscriminació permet les accions populars plantejades per ONG que representin l'interès públic quan la discriminació es basa en una característica protegida que és una característica essencial de la persona i afecta a un grup més gran de víctimes potencials que no poden ser identificades¹⁶⁰⁴. Amb el benentès que sota certs ordenaments jurídics d'estats membres de la UE l'acció popular es duu a terme quan les persones demandants individuals no són identificades, i únicament cal que l'organització demostrï un interès en la protecció del dret a la igualtat de tractament¹⁶⁰⁵, des de l'any 2000, davant del TJUE s'han dut a terme diverses accions populars¹⁶⁰⁶, com per exemple el cas Feryn¹⁶⁰⁷, Age Concern England¹⁶⁰⁸, Test-Achats and Others¹⁶⁰⁹ i Associația Accept¹⁶¹⁰. Tot i així, cal remarcar que les accions populars són diferents tècnicament a les accions col·lectives (ja sigui per protegir interessos col·lectius, difusos o (pluri)individuals homogenis), ja que aquestes últimes s'inicien per part d'organitzacions o organismes públics en benefici dels interessos d'un grup de persones, i les accions populars per perseguir l'interès públic de manera abstracta¹⁶¹¹. La qüestió interpretativa resideix en què en aquests estats membres s'entén l'interès general de manera abstracta, com l'interès per protegir un grup particular de persones en contra de violacions sistèmiques de drets que representen un interès públic en una societat definida com a democràtica, i en canvi com hem vist anteriorment, també es pot entendre com un mer interès en la legalitat, sense cap qualificació subjectiva especial¹⁶¹².

¹⁶⁰² ORGANISATION FOR SECURITY AND CO-OPERATION IN EUROPE (OSCE), *Use of actio popularis...*, *op. cit.*, November 2016, pàg. 16.

¹⁶⁰³ *Idem*, pàg. 12-14.

¹⁶⁰⁴ FARKAS, L., "Limited enforcement possibilities under European anti-discrimination...", *op. cit.*, 2010, pàg. 186.

¹⁶⁰⁵ FARKAS, L., "Collective actions under European anti-discrimination law", *op. cit.*, 2014, pàg. 31.

¹⁶⁰⁶ *Idem*, pàg. 35.

¹⁶⁰⁷ TJCE (Sala Segona), sentència de 10 de juliol de 2008, C-54/07, Centrum voor gelijkheid van kansen en voor racismebestrijding contra Firma Feryn NV.

¹⁶⁰⁸ TJUE (Sala Tercera), sentència de 5 de març de 2009, C-388/07, The Queen, on the application of The Incorporated Trustees of the National Council for Ageing (Age Concern England) contra Secretary of State for Business, Enterprise and Regulatory Reform.

¹⁶⁰⁹ TJUE (Gran Sala), sentència d'1 de març de 2011, C-236/09, Association Belge des Consommateurs Test-Achats ASBL y otros contra Conseil des ministres.

¹⁶¹⁰ TJUE (Sala Tercera), sentència de 25 d'abril de 2013, C-81/12, Asociația Accept contra Consiliul Național pentru Combaterea Discriminării.

¹⁶¹¹ BENEDI LAHUERTA, S., "Enforcing EU Equality Law...", *op. cit.*, 2018, pàg. 788.

¹⁶¹² COROMINAS BACH, S., *La legitimación activa en las acciones col·lectivas*, *op. cit.*, 2015, pàg. 38.

La UE ha dut a terme accions institucionals per millorar la situació col·lectiva d'alguns grups socials afectats per la discriminació, però aquesta acció no ha anat acompanyada de la promoció del desenvolupament per part dels estats membres de mecanismes de reparació col·lectiva¹⁶¹³. Cada estat membre ha adoptat un conjunt diferent de processos i salvaguardes en relació amb la litigació col·lectiva, basats simplement en el resultat dels debats polítics nacionals¹⁶¹⁴. A escala europea, interessaria una norma que obligués a tots els estats membres a proveir mecanismes per tal que les organitzacions poguessin protegir els interessos difusos, sense representar una víctima individual identificada, i no haver de confiar que l'estat membre hagi creat una norma en la qual ho faci possible. De la mateixa manera, seria necessari clarificar el contingut tant de l'acció de cessació com de l'acció reparadora a escala europea. La Directiva 2020/1828 hagués estat una bona oportunitat, però el seu àmbit d'aplicació és limitat. Com hem vist, seguint les exigències de les directives antidiscriminació no hi ha una cobertura tan exhaustiva com la de la Directiva 2020/1828 en relació a com s'ha de configurar un mecanisme d'acció col·lectiva, i era una bona oportunitat per crear-lo. En tot cas, a Espanya, no obstant que s'hagi incorporat la defensa d'interessos difusos en l'àmbit antidiscriminació -no requerit en les directives antidiscriminació-, la regulació de l'acció col·lectiva en l'àmbit de la discriminació és dispersa i confusa, com veurem seguidament.

2.2. Acció de representació a Espanya en matèria de discriminació algorítmica

Com ja hem dit, no hi ha un procediment d'acció col·lectiva comú en la UE en matèria antidiscriminació (únicament la Directiva 2020/1828 que és un instrument de mínims - per tal que els estats membres puguin tenir si més no un mecanisme d'acció col·lectiva en matèria de dret de les persones consumidores¹⁶¹⁵-). Per consegüent, ens remetem a les

¹⁶¹³ BENEDI LAHUERTA, S., "Enforcing EU Equality Law...", *op. cit.*, 2018, pàg. 788.

¹⁶¹⁴ HODGES, C., "Collective Redress: The Need for New Technologies", *Journal of Consumer Policy*, vol. 42, núm. 1, 2019, pàg. 62.

¹⁶¹⁵ Malgrat que la Directiva 2020/1828, en la seva regulació de mínims, no integra l'obligació específica en relació amb les situacions de discriminació en les relacions de consum, la normativa espanyola sí que ho contempla. L'article 47.1 del Real Decreto Legislativo 1/2007, de 16 de noviembre, por el que se aprueba el texto refundido de la Ley General para la Defensa de los Consumidores y Usuarios y otras leyes complementarias (RDLETRLGDCU), estableix que són infraccions en matèria de defensa de les persones consumidores i usuàries tota actuació discriminatòria contra persones consumidores vulnerables independentment del motiu o contra qualsevol persona consumidora o usuària per l'exercici dels drets, ja

directives antidiscriminació, que requereixen a cada estat membre de la UE -sense entrar a valorar com s'han d'aconseguir els objectius que exposen- disposar d'un mecanisme que legítimi certes entitats habilitades pel dret nacional a iniciar en nom de la persona demandant, o en el seu suport i la seva autorització, qualsevol procediment judicial o administratiu, sense anar més enllà. L'ordenament jurídic espanyol permet la intervenció de persones jurídiques per a la defensa d'interessos col·lectius, reconeixent en l'article 7.3 de la Ley Orgánica 6/1985, de 1 de julio, del Poder Judicial (d'ara en endavant, LOPJ), la legitimació de corporacions, associacions i grups que resultin afectats o que estiguin legalment habilitats per la seva defensa i promoció¹⁶¹⁶.

En primer lloc, per via de la representació voluntària (que es produeix, no per imperatiu legal, sinó per la pròpia voluntat de la persona representada) i no de legitimació extraordinària, es reconeix l'actuació processal de certes persones jurídiques¹⁶¹⁷, com en el Real Decreto Legislativo 1/2013, de 29 de noviembre, por el que se aprueba el Texto Refundido de la Ley General de derechos de las personas con discapacidad y de su inclusión social, que disposa en el seu article 76 que, addicionalment a la legitimació individual, les persones jurídiques legalment habilitades per a la defensa dels drets i interessos legítims col·lectius podran actuar en un procés en nom i interès de les persones que així ho autoritzin, amb la finalitat de fer efectiu el dret a la igualtat d'oportunitats defensant els drets individuals i recaient en aquestes persones els efectes d'aquella autorització¹⁶¹⁸. També, l'article 31 de la Llei 62/2003, estableix que les persones jurídiques que estiguin legalment habilitades per a la defensa dels drets i interessos legítims col·lectius podran actuar en processos judicials en nom de la persona demandant que així ho autoritzi amb la finalitat de fer efectiu el principi d'igualtat de tracte de les persones pel seu origen racial o ètnic¹⁶¹⁹.

En segon lloc, en la normativa antidiscriminació espanyola, actualment, en relació amb la legitimació, es compta amb l'article 19.1.i de la Ley 29/1998, de 13 de julio, reguladora

sigui no atenent a les seves demandes, negant-los l'accés als establiments o donant-los un tracte o imposant unes condicions desiguals i les conductes discriminatòries en l'accés a béns i a la prestació de serveis i, en especial, les previstes com a tals en la Llei Orgánica 3/2007, quan no siguin constitutives de delictes.

¹⁶¹⁶ Article 7.3 de la LOPJ.

¹⁶¹⁷ MORENO CATENA, V., *Derecho Procesal Civil. Parte general*, op. cit., 2021, pàg. 100.

¹⁶¹⁸ Article 76 del Real Decreto Legislativo 1/2013, de 29 de noviembre, por el que se aprueba el Texto Refundido de la Ley General de derechos de las personas con discapacidad y de su inclusión social.

¹⁶¹⁹ Article 31 de la Llei 62/2003.

de la Jurisdicció Contencioso-administrativa¹⁶²⁰ (d'ara en endavant, Llei 29/1998), introduït per la Llei 15/2022, i l'article 11 bis de la LEC, que va ser introduït per la Llei Orgànica 3/2007, en la primera regulació de la tutela col·lectiva en la matèria¹⁶²¹ i modificat per la Llei 15/2022, ampliant la possibilitat de tutela col·lectiva a tots els motius de discriminació coberts per la legislació espanyola, no només la igualtat efectiva entre homes i dones¹⁶²². Tant l'article 19.1.i de Llei 29/1998 com l'article 11 bis de la LEC legitimen, en el primer apartat, per a la defensa del dret a la igualtat de tracte i no discriminació, a més de les persones afectades i sempre amb la seva autorització, a l'Autoritat Independent per a la Igualtat de Tracte i la No Discriminació (d'ara en endavant, Autoritat Independent), en relació a les persones afiliades i associades als mateixos a partits polítics, sindicats, associacions professionals de treballadors autònoms, organitzacions de persones consumidores i usuàries i associacions i organitzacions legalment constituïdes que tinguin entre les seves finalitats la defensa i promoció dels drets humans¹⁶²³ (podríem dir que aquest és l'article que respon a la transposició de les

¹⁶²⁰ Ley 29/1998, de 13 de julio, reguladora de la Jurisdicción Contencioso-administrativa. BOE, núm. 167, de 14 de julio de 1998. España.

¹⁶²¹ L'article 11 bis de la LEC abans de ser modificat per la Llei 15/2022 versava de la següent manera: "1. Para la defensa del derecho de igualdad de trato entre mujeres y hombres, además de los afectados y siempre con su autorización, estarán también legitimados los sindicatos y las asociaciones legalmente constituidas cuyo fin primordial sea la defensa de la igualdad de trato entre mujeres y hombres, respecto de sus afiliados y asociados, respectivamente. 2. Cuando los afectados sean una pluralidad de personas indeterminada o de difícil determinación, la legitimación para demandar en juicio la defensa de estos intereses difusos corresponderá exclusivamente a los organismos públicos con competencia en la materia, a los sindicatos más representativos y a las asociaciones de ámbito estatal cuyo fin primordial sea la igualdad entre mujeres y hombres, sin perjuicio, si los afectados estuvieran determinados, de su propia legitimación procesal. 3. La persona acosada será la única legitimada en los litigios sobre acoso sexual y acoso por razón de sexo."

¹⁶²² La llei 19/2020, en el seu article 26.2 també estableix que "Sens perjudici de la legitimació individual de la persona afectada, els sindicats, les associacions professionals de treballadors autònoms, les organitzacions de consumidors i usuaris i les associacions i organitzacions constituïdes legalment que tinguin entre llurs fins la defensa i la promoció dels drets humans o dels col·lectius i els drets a què fa referència aquesta llei, i compleixin els requisits de l'apartat 4, poden tenir la consideració de persona interessada en els procediments administratius en què l'Administració s'hagi de pronunciar amb relació a una situació de discriminació, sempre que comptin amb l'autorització de la persona o persones afectades. Si la persona o les persones afectades són una pluralitat indeterminada o de difícil determinació, aquesta autorització no és necessària, sens perjudici que els qui es considerin afectats puguin participar en el procediment." Tanmateix en cap cas estableix que una institució pública pugui considerar-se persona interessada.

¹⁶²³ Article 11 bis.1 de la LEC. Cal puntualitzar que l'article 11 bis no legitima al Ministeri Fiscal a dur a terme aquestes accions, a diferència de l'article 11.5 de la LEC que sí que el legitima per exercitar qualsevol acció en defensa de les persones consumidores i usuàries; Article 19.1.i de la Llei 29/1998, reguladora de la jurisdicció Contencioso-administrativa. En relació amb el procediment administratiu, cal fer referència a l'article 31.2 de la Llei 62/2003, que estableix que "a los efectos de lo establecido en el artículo 4.2 de la Ley 39/2015, de 1 de octubre, del Procedimiento Administrativo Común de las Administraciones Públicas, los sindicatos, las asociaciones profesionales de trabajadores autónomos, las organizaciones de personas consumidoras y usuarias y las asociaciones y organizaciones legalmente constituidas que tengan entre sus fines la defensa y promoción de los derechos humanos y cumplan los requisitos fijados en el artículo 29 de la presente ley, podrán tener la consideración de interesado en los procedimientos administrativos en los

directives antidiscriminació). En el segon apartat, quan es tracti de persones afectades que constitueixin una pluralitat indeterminada o de difícil determinació (és a dir, per protegir interessos difusos), es legitima a l'Autoritat Independent, partits polítics, sindicats, associacions professionals de treballadors autònoms més representatives, organitzacions de persones consumidores i usuaris d'àmbit estatal, organitzacions, d'àmbit estatal o de l'àmbit territorial on es produeixi la situació de discriminació que tinguin entre les seves finalitats la defensa i promoció dels drets humans, sense perjudici en tot cas de la legitimitat individual d'aquelles persones afectades que estiguessin determinades¹⁶²⁴ (podríem dir que en aquest apartat es va més enllà d'allò requerit en les directives antidiscriminació).

En referència exclusivament a la LEC, l'article 11 bis 1 fa necessària una autorització de les persones afectades per tal que les entitats legitimades puguin demanar tutela dels seus drets individuals sense que l'autorització pugui ser obtinguda pels procediments propis del règim general d'aquesta matèria. El règim general per constatar la voluntat de les persones que poden ser individualment afectades per la tutela judicial sol·licitada col·lectivament l'estableix l'article 15 ter de la LEC¹⁶²⁵, en relació amb els articles 221.1, 222.3 i 519 de la mateixa norma¹⁶²⁶. Es preveu un canvi en aquestes normes quan s'aprovi la Ley de acciones de representación para la protección de los intereses colectivos de los consumidores que transposa la Directiva 2020/1828 a Espanya.

La regulació processal a Espanya de les accions col·lectives ha experimentat canvis rellevants, s'han reformat preceptes de la matèria tant en la LEC com en altres lleis que tradicionalment les han contemplat -lleis de protecció de consumidors i usuaris, generals

que la Administración tenga que pronunciarse en relación con una situación de discriminación prevista en esta ley, siempre que cuenten con la autorización de la persona o personas afectadas. No será necesaria esta autorización cuando las personas afectadas sean una pluralidad indeterminada o de difícil determinación, sin perjuicio de que quienes se consideren afectados puedan también participar en el procedimiento”.

¹⁶²⁴ Article 11 bis.2 LEC; Article 19.1.i de la Llei 29/1998. Aquest article estableix, a més, que la persona assetjada serà l'única legitimada en els casos de litigis sobre assetjament sexual i assetjament discriminatori.

¹⁶²⁵ Article 15 ter de la LEC. Aquest article va ser introduït per la Llei 15/2022 i disposa que: “1. En los procesos promovidos por la Autoridad Independiente para la Igualdad de Trato y la No Discriminación, los partidos políticos, sindicatos, asociaciones profesionales de trabajadores autónomos, organizaciones de personas consumidoras y usuarias y asociaciones y organizaciones legalmente constituidas, que tengan entre sus fines la defensa y promoción de los derechos humanos, se llamará al proceso a quienes tengan la condición de personas afectadas por haber sufrido la situación de discriminación que dio origen al proceso, para que hagan valer su derecho o interés individual. 2. El órgano judicial que conozca de alguno de estos procesos comunicará su iniciación al Ministerio Fiscal para que, de conformidad con las funciones que le son propias, valore la posibilidad de su personación.”

¹⁶²⁶ ORTELLS RAMOS, M., “Tutela judicial civil colectiva y nuevos modelos de los servicios de defensa jurídica en España”, *Revista General de Derecho Procesal*, vol. 48, 2019, pàg. 12.

i sectorials, de competència, publicitat i condicions generals de la contractació- i lleis noves han introduït normes en àrees del dret material que abans havien restat alienes a la regulació de les accions col·lectives -igualtat efectiva entre homes i dones, igualtat d'oportunitats de persones amb discapacitat¹⁶²⁷, com també en l'àmbit de la no discriminació en general.

Tanmateix, l'àmbit d'aplicació de la tutela col·lectiva està acotat a la tutela judicial en sectors específics del dret privat i la més àmplia aplicació correspon als drets i interessos jurídics de persones consumidores i usuàries. La pròpia LEC denota aquesta preponderància ja que les disposicions sobre la matèria que s'introdueixen en l'estructura sistemàtica del seu text només mencionen la tutela de drets de persones consumidores i usuàries¹⁶²⁸. Hi ha autors que consideren que aquesta limitació a les altres matèries (com per exemple el medi ambient, o afegim nosaltres, la no discriminació) és injustificada i proposen que sigui superada ja sigui a través d'analogia o a través de reforma legal¹⁶²⁹ (essent aquesta última la que ha obert la via a que casos de discriminació recaiguin en la tutela col·lectiva¹⁶³⁰). De tota manera, la regulació de l'acció col·lectiva es troba dispersa en la LEC¹⁶³¹ i s'espera que a través de la transposició de la Directiva 2020/1828 s'uneixin

¹⁶²⁷ *Idem*, pàg. 5.

¹⁶²⁸ *Idem*, pàg. 8.

¹⁶²⁹ Vegeu: GARNICA MARTÍN, J. F., “Artículo 11. Legitimación para la defensa de derechos e intereses de consumidores y usuarios”, a FERNÁNDEZ-BALLESTEROS LÓPEZ, M.A., RIFÁ SOLER, J.M., i VALLS GOMBAU, J.F. (coords.), *Comentarios a la nueva Ley de Enjuiciamiento Civil*, Tomo I, Barcelona, Atelier, 2001, pàg. 164 i 165.

¹⁶³⁰ ORTELLS RAMOS, M., “Tutela judicial civil colectiva y nuevos modelos...”, *op. cit.*, 2019, pàg. 9.

¹⁶³¹ NIETO MORALES, C. i ABASCAL MONEDERO, P. J., *Reclamaciones en Materia de Consumo*, Madrid, Dykinson, 2016, pàg. 213-215. El procediment d'acció col·lectiva es desenvolupa en els següents articles de la LEC: l'article 6, que preveu quines persones estan capacitades per ser part en el procés civil, entre elles grups de persones consumidores i usuàries com entitats habilitades conforme la normativa europea; l'article 11, que especifica qui ostenta la legitimació per a la defensa de drets i interessos de persones consumidores i usuàries; l'article 13.1, que preveu el dret de persones consumidores i usuàries a intervenir en els processos iniciats per entitats en defensa d'interessos col·lectius; l'article 15.1, que preveu publicitat i intervenció de les persones perjudicades en processos promoguts per associacions que preveu l'ús de mitjans de comunicació de l'àmbit territorial de les persones que es vegin afectades per cridar-les a participar; l'article 15.2, que disposa que quan estiguin determinades les persones afectades es precisa comunicació prèvia a la presentació de la demanda, i la persona afectada podrà intervenir en qualsevol moment del procés però sense que es retrotraguin les actuacions; l'article 15.3, que preveu que quan siguin indeterminades, la crida a persones perjudicades suspèn el curs del procediment per, com a màxim, dos mesos, sense que amb posterioritat puguin personar-se, sense perjudici que facin valer els seus drets dels articles 221 -efectes de la sentència-, i 519 -acció executiva basada en la sentència condemnatòria-; l'article 52.1.16, que estableix un fur especial, senyalant que en les accions de cessació en defensa d'interessos tant difusos com col·lectius, serà competent el Tribunal de l'indret on estigui el domicili de la persona demandada, llevat que no en tingui, llavors serà el seu domicili, llevat que aquest domicili no es trobi en l'estat espanyol, llavors serà el domicili de la part actora; l'article 78.4, que preveu l'acumulació de processos d'aquesta naturalesa, inclús d'ofici, d'acord amb les regles generals de la llei; l'article 221, que es refereix als requisits interns que ha de contenir la sentència dictada en processos de consum, fixant les normes de determinació i reconeixement de la condició de beneficiaris a les persones afectades sense que

i aclareixin les disposicions en matèria d'accions col·lectives en l'àmbit del consum (que es podrien aplicar per analogia a l'acció col·lectiva en matèria de discriminació) i es posi remei a la dispersió de la regulació col·lectiva que es troba al llarg de la LEC¹⁶³², amb l'expectativa d'un nou procés especial a la LEC i no un procés ordinari amb especialitats.

Aquesta dispersió no la trobem, per exemple, a França, on la Loi 2016-1547, de 18 de novembre de 2016, de modernisation de la justice du XXIe siècle¹⁶³³, en els seus articles 86 a 88, va incorporar una acció col·lectiva (en francès *action de groupe*) en matèria de discriminació. Arran de l'article 86, es va modificar l'article 10 de la Loi n° 2008-496 du 27 mai 2008 portant diverses dispositions d'adaptation au droit communautaire dans le domaine de la lutte contre les discriminations¹⁶³⁴, possibilitant que una associació declarada regularment durant cinc anys -com a mínim-, reconeguda en la lluita contra les discriminacions o en l'àmbit de la discapacitat, dugués a terme una acció col·lectiva tant de cessació com reparadora quan moltes persones físiques haguessin patit discriminació directa o indirecta fonamentada en un mateix motiu imputable a una mateixa persona¹⁶³⁵. És en el Décret n° 2017-888 du 6 mai 2017 relatif à l'action de groupe et à l'action en reconnaissance de droits prévues aux titres V et VI de la loi n° 2016-1547 du 18 novembre 2016 de modernisation de la justice du XXIe siècle, on es desenvolupa l'exercici de

es prevegi cap actuació de caràcter processal; els articles 249.4 i 5 i 250.12, que estableixen que aquesta tipologia de processos es tramiten en tot cas per Judici Verbal; l'article 256.1.6, que estableix una diligència preliminar específica destinada a la concreció dels integrants del grup d'afectats quan, sense estar determinats, siguin fàcilment determinables, senyalant-se una norma especial de competència de l'article 257; l'article 519, que estableix l'acció executiva fundada en la sentència de condemna sense determinació individual de beneficiats, introduint el tràmit precís per a tal reconeixement, que parteix de la necessària sol·licitud d'una o diverses persones afectades, amb audiència de la persona condemnada, a l'objecte de resoldre, d'acord amb les dades característiques i requisits establerts en la sentència, sobre la procedència d'expedir auto per iniciar l'execució; i l'article 728, que estableix la previsió expressa de prestar fiança per la part sol·licitant d'una mesura cautelar, i en el seu tercer apartat, últim paràgraf, precisa que el Tribunal podrà dispensar de prestació de fiança a la persona sol·licitant en els casos de defensa d'interessos col·lectius i difusos de persones consumidores i usuàries ateses les circumstàncies del cas, així com l'entitat econòmica i la repercussió social dels diferents interessos afectats.

¹⁶³² CHOCHRÓN GIRÁLDEZ, A. M., "El marco jurídico de la Directiva 2020/1828...", *op. cit.*, 2022, pàg. 286.

¹⁶³³ Loi 2016-1547, du 18 novembre 2016, de modernisation de la justice du XXIe siècle. JORF n° 0269 du 19 novembre 2016. France.

¹⁶³⁴ Article 10 de la Loi n° 2008-496 du 27 mai 2008 portant diverses dispositions d'adaptation au droit communautaire dans le domaine de la lutte contre les discriminations. JORF n°0123 du 28 mai 2008. France.

¹⁶³⁵ Tanmateix, malgrat que a França existeixi l'acció de grup en matèria de discriminació, conté moltes limitacions que són subratllades per la Defensoria de Drets (l'organisme nacional d'igualtat francès). Vegeu: DÉFENSEUR DES DROITS, *Avis du Défenseur des droits n°23-03*, Paris, 23 février 2023. Per superar les limitacions proposa obrir la legitimació per dur a terme accions col·lectives en matèria de discriminació (reservada actualment a organitzacions sindicals i a tota associació regularment declarada després de cinc anys que intervingui en la lluita envers la discriminació o que operi dins l'àmbit de la discapacitat), com també crear un fons per assegurar el finançament d'accions col·lectives.

l'acció¹⁶³⁶.

En relació amb la prova, la Llei 15/2022 modifica tant l'apartat 5 de l'article 217 de la LEC, com l'apartat 7 de l'article 60 de la Ley 29/1998¹⁶³⁷, i introdueix l'article 77.3 bis en la Ley 39/2015, de 1 de octubre, del Procedimiento Administrativo Común de les Administraciones Públicas¹⁶³⁸ (d'ara en endavant, Llei 39/2015). En els tres articles s'estableix, de la mateixa manera, que quan la persona interessada al·legui discriminació i aporti indicis fundats sobre la seva existència, correspondrà a la persona a qui s'imputi la situació discriminatòria l'aportació d'una justificació objectiva i raonable, suficientment provada, de les mesures adoptades i de la seva proporcionalitat, possibilitant que l'òrgan administratiu (en la Llei 39/2015) i l'òrgan judicial, d'ofici o a instància de part (en la LEC i la Ley 29/1998) pugui recavar informe dels organismes públics competents en matèria d'igualtat¹⁶³⁹. A través d'aquests articles, les exigències en relació amb el repartiment de la càrrega de la prova, que ja s'anunciaven en les directives antidiscriminació¹⁶⁴⁰, s'estenen a qualsevol motiu de discriminació.

En relació amb la indemnització, la Llei 15/2022 disposa en el seu article 27 que la persona física o jurídica que causi discriminació per algun dels motius previstos en la llei repararà el dany causat proporcionant una indemnització i restituint la víctima a la situació anterior a l'incident discriminatori quan sigui possible¹⁶⁴¹. També especifica que acreditada la discriminació es presumirà dany moral, que es valorarà atenent les circumstàncies del cas, la concurrència o interacció de diverses causes de discriminació previstes per la llei i a la gravetat de la lesió efectivament produïda¹⁶⁴². L'article 28 de la mateixa normativa estableix, per la seva banda, que la tutela judicial enfront les vulneracions del dret a la igualtat de tracte i no discriminació comprendrà l'adopció de les mesures necessàries per posar fi a la discriminació de la qual es tracti, i en particular, al cessament immediat de la discriminació, podent acordar l'adopció de mesures cautelars

¹⁶³⁶ Décret n° 2017-888 du 6 mai 2017 relatif à l'action de groupe et à l'action en reconnaissance de droits prévues aux titres V et VI de la loi n° 2016-1547 du 18 novembre 2016 de modernisation de la justice du XXIe siècle. JORF n°0109 du 10 mai 2017. France.

¹⁶³⁷ Article 60.7 de la Llei 29/1998.

¹⁶³⁸ Ley 39/2015, de 1 de octubre, del Procedimiento Administrativo Común de las Administraciones Públicas. BOE, núm. 236, de 02 d'octubre de 2015. Espanya.

¹⁶³⁹ Article 217.5 de la LEC; article 60.7 de la Ley 29/1998; i article 77.3.bis de la Ley 39/2015.

¹⁶⁴⁰ Article 8.1 de la Directiva 2000/43/CE; article 10.1 de la Directiva 2000/78/CE, article 9.1 de la Directiva 2004/113/CE; article 19.2 de la Directiva 2006/54/CE.

¹⁶⁴¹ Article 27.1 de la Llei 15/2022.

¹⁶⁴² Article 27.1 de la Llei 15/2022.

dirigides a la prevenció de violacions imminents o ulteriors, la indemnització dels danys i perjudicis causats i el restabliment de la persona perjudicada en el ple exercici del seu dret, amb independència de la seva nacionalitat, de si és major o menor d'edat o de si gaudeix o no de residència legal¹⁶⁴³.

Havent aclarit quines normes regulen les accions col·lectives en matèria de discriminació tant a escala europea com espanyola, seguidament abordarem el valor afegit que pot aportar l'acció col·lectiva en la discriminació algorítmica.

2.3. Avantatges de l'acció de representació en matèria de discriminació algorítmica en comparació amb l'acció individual

Com ja hem vist en l'apartat 3 i 4 del capítol IV, la sentència CHEZ va reiterar la idea fonamental, ja avançada amb caràcter general en la sentència Feryn, que l'existència d'una discriminació no exigeix que la víctima de la conducta discriminatòria hagi de ser singular i identificable, i tant en la sentència Accept com en la Rete Lenford, el TJUE va admetre la suficiència d'un dany hipotètic per establir una discriminació *prima facie*. En comptes de focalitzar-se en injustícies patides per part de víctimes individuals es va focalitzar en patrons, tendències i escenaris de discriminació¹⁶⁴⁴.

Aquesta jurisprudència és positiva en relació amb la discriminació algorítmica, en la qual la persona perjudicada molt probablement no és conscient que està patint discriminació. Més que en relació amb una acció reparadora, sembla pertinent veure les possibilitats d'una acció de cessació, que aconseguixi que es deixi de dur a terme una acció que causa discriminació envers certs col·lectius determinats de població. Pren rellevància aquí, que la litigació individual només implicaria la persona que s'hagués vist discriminada (i hagués detectat aquesta discriminació). En canvi, en el cas de la discriminació algorítmica, és més evident que és el col·lectiu qui constantment es pot veure perjudicat pel sistema, per tant és necessari litigar en el sentit que la sentència pugui tenir una incidència més àmplia que no únicament individual. Conseqüentment, la possibilitat de

¹⁶⁴³ Article 28 de la Llei 15/2022.

¹⁶⁴⁴ FARKAS, L., "Limited enforcement possibilities under European anti-discrimination...", *op. cit.*, 2010, pàg. 186.

defensar interessos col·lectius i difusos per part d'entitats legitimades en comptes de que recaigui el pes de l'acció només en la persona perjudicada, millora les possibilitats de lluita envers la discriminació algorítmica. Tanmateix, depèn dels estats membres que es contempli la possibilitat de defensar interessos difusos ja que les directives antidiscriminació no ho exigeixen.

Quant a la prova, tal com hem dit anteriorment, en el cas de tractar-se d'una acció col·lectiva de discriminació per protegir interessos (pluri)individuals homogenis, la posició de diverses persones unides en una única acció per fer front a la lesió (discriminació arran d'un sistema algorítmic) causada per un fet comú (decisió presa mitjançant un sistema algorítmic) millora la seva posició en comparació amb l'acció duta a terme individualment, ja que el fet que hi hagi diverses persones que comparteixin una característica protegida que s'hagin vist lesionades per una decisió presa per un sistema algorítmic, ja pot ser per si mateix un indicatiu de discriminació. A la vegada, si es duu a terme una acció per defensar interessos (pluri)individuals homogenis, en el cas de la discriminació, com també s'ha esmentat anteriorment, també es veuen implicats interessos col·lectius, atès que la decisió és discriminatòria com a conseqüència d'un tracte menys favorable del sistema a un grup concret de persones.

Cal fer palès, en relació amb l'obtenció de prova de discriminació, que la normativa antidiscriminació no dona eines per obtenir indicis de discriminació algorítmica. Si bé és cert que posteriorment a la presentació de prova hi haurà una inversió de la càrrega de la mateixa, per tal d'obtenir indicis no hi ha cap suport¹⁶⁴⁵. Cobra nova rellevància l'article 4.b de la Propuesta de Directiva del Consejo relativa a la carga de la prueba en el ámbito de la igualdad de retribución y trato entre hombres y mujeres segons el qual les dues parts podrien disposar de qualsevol informació pertinent en poder de l'altra part per fer valer els seus drets, sempre que la divulgació de la informació no afectés greument als interessos de l'altra part per motius aliens al litigi¹⁶⁴⁶ (vegeu apartat 4.2 del capítol IV).

El que pot aportar l'acció col·lectiva des de la normativa antidiscriminació, principalment, és proporcionar l'existència d'algú altre, a part de les persones perjudicades, que pugui lluitar envers la discriminació algorítmica, atesa la situació de

¹⁶⁴⁵ Article 8.1 de la Directiva 2000/43/CE; article 10.1 de la Directiva 2000/78/CE, article 9.1 de la Directiva 2004/113/CE; article 19.2 de la Directiva 2006/54/CE.

¹⁶⁴⁶ COMISIÓN DE LAS COMUNIDADES EUROPEAS, *Propuesta de Directiva del Consejo relativa a la carga de la prueba en el ámbito de la igualdad...*, op. cit., 27 de mayo de 1988, art. 4.b.

desempament jurídic en què es troben les persones que s'han vist perjudicades mitjançant un sistema algorítmic (vegeu capítol VIII). Ja va plantejar Xenidis que el fet que s'involucrin actors col·lectius podria disminuir la càrrega de la prova en casos de discriminació algorítmica¹⁶⁴⁷. A més, en l'ús de sistemes algorítmics, la protecció col·lectiva -en comparació amb la protecció individual-, esdevé més rellevant, atès que si no es demana el cessament de l'ús del sistema algorítmic, el col·lectiu al qual pertany la persona perjudicada (dones, persones d'un grup ètnic determinat, d'una edat determinada, d'una orientació sexual determinada, etc.) seguirà veient-se perjudicat.

Resumidament, la normativa antidiscriminació europea no obliga els estats membres a disposar d'una acció col·lectiva per protegir interessos difusos, sinó que les directives antidiscriminació únicament obliguen a l'establiment de la possibilitat que associacions, organitzacions i entitats jurídiques puguin iniciar en nom de la persona demandant o en el seu suport i autorització procediments judicials o administratius. A Espanya, és possible la defensa d'interessos difusos de persones que s'han vist discriminades, i no és únicament la persona discriminada la que ha de vetllar per la seva protecció, sinó que entitats amb més accés a informació i recursos, per exemple els organismes d'igualtat nacionals (com veurem en el capítol VIII), poden desafiar decisions que causin discriminació algorítmica i plantejar la cessació del sistema algorítmic que causa la discriminació. Tot seguit abordem l'ús de l'acció col·lectiva en matèria de protecció de dades per abordar infraccions de discriminació algorítmica.

3. Normativa de protecció de dades i l'acció de representació en matèria de discriminació algorítmica

En aquest apartat estudiem quines normes tant europees (3.1), com espanyoles (3.2) habiliten la possibilitat d'iniciar una acció col·lectiva en matèria de protecció de dades en el context de la discriminació algorítmica. Finalment, analitzem el valor afegit que poden aportar les accions col·lectives de protecció de dades en la lluita envers la discriminació

¹⁶⁴⁷ XENIDIS, R., "Turning EU equality law to algorithmic discrimination...", *op. cit.*, 2020, pàg. 750; TJUE (Gran Sala), sentència de 23 d'abril de 2020, C 507/18, NH contra Associazione Avvocatura per i diritti LGBTI - Rete Lenford, par. 65; TJUE, sentència de 10 de juliol de 2008, C-54/07, Centrum voor gelijkheid van kansen en voor racismebestrijding contra Firma Feryn NV, par. 27.

algorítmica (3.3).

3.1. Acció de representació de protecció de dades a la UE en el context de la discriminació algorítmica

En l'apartat 5 del capítol V d'aquest treball hem analitzat com l'RGPD proporciona en el seu article 80 la possibilitat per part de la persona interessada d'encomanar a una entitat, organització o associació de presentar en el seu nom una reclamació davant l'autoritat de control o òrgan jurisdiccional i l'exercici del seu dret a ser indemnitzada¹⁶⁴⁸ (que estableix l'article 82 de l'RGPD¹⁶⁴⁹), i la possibilitat que els estats membres disposin que una entitat, organització o associació amb independència del mandat de la persona interessada, en el cas d'una infracció de l'RGPD, tingui dret a presentar una reclamació davant l'autoritat de control competent (art. 77 RGPD)¹⁶⁵⁰, a la tutela judicial efectiva contra una autoritat de control (art. 78 RGPD)¹⁶⁵¹ i a la tutela judicial efectiva contra una persona responsable o encarregada del tractament (art. 79 RGPD)¹⁶⁵². En la sentència de 28 d'abril de 2022, *Meta Platforms Ireland Limited*, el TJUE va avalar que les associacions de persones consumidores exercitin accions, sense mandat (no necessitant

¹⁶⁴⁸ Article 80.1 de l'RGPD.

¹⁶⁴⁹ L'article 82 de l'RGPD estableix que tota persona que hagi patit danys i perjudicis materials o immaterials a conseqüència d'una infracció de l'RGPD podrà rebre per part de la persona responsable de tractament una indemnització per danys i perjudicis.

¹⁶⁵⁰ Per una anàlisi amb profunditat de l'article 77 de l'RGPD, vegeu: VILLALBA CANO, L., "El derecho a presentar una reclamación ante una autoridad de control (Comentario al artículo 77 RGPD)", a TRONCOSO REIGADA, A., (dir.) i VILLALBA CANO, L., (coord.), *Comentario al Reglamento General de Protección de Datos y a la Ley Orgánica de Protección de Datos Personales y Garantía de los Derechos Digitales*, Tomo II, Pamplona, Aranzadi, 2021, pàg. 2967-2987.

¹⁶⁵¹ Per una anàlisi amb profunditat de l'article 78 de l'RGPD, vegeu: VILLALBA CANO, L., "El derecho a la tutela judicial efectiva contra una autoridad de control (Comentario al artículo 78 RGPD)", a TRONCOSO REIGADA, A., (dir.) i VILLALBA CANO, L., (coord.), *Comentario al Reglamento General de Protección de Datos y a la Ley Orgánica de Protección de Datos Personales y Garantía de los Derechos Digitales*, Tomo II, Pamplona, Aranzadi, 2021, pàg. 2989-3005.

¹⁶⁵² Per una anàlisi amb profunditat de l'article 79 de l'RGPD, vegeu: VILLALBA CANO, L., "El derecho a la tutela judicial efectiva contra un responsable o encargado del tratamiento (Comentario al artículo 79 RGPD)", a TRONCOSO REIGADA, A., (dir.) i VILLALBA CANO, L., (coord.), *Comentario al Reglamento General de Protección de Datos y a la Ley Orgánica de Protección de Datos Personales y Garantía de los Derechos Digitales*, Tomo II, Pamplona, Aranzadi, 2021, pàg. 3007-3018. L'autora precisa que en virtut de l'article 79.1 de l'RGPD, qualsevol persona que consideri que ha vist vulnerats els seus drets podrà sol·licitar la tutela judicial efectiva contra la persona responsable o la persona encarregada del tractament que ha ocasionat la vulneració, essent la via judicial l'única opció per obtenir la indemnització de l'article 82 per danys i perjudicis materials e immaterials produïts per l'incompliment de l'RGPD (pàg. 3010 i 3011).

una víctima identificada), i amb independència de la vulneració de drets concrets (possibilitat d'un dany hipotètic) de la persona interessada contra qui infringís la normativa en matèria de protecció de dades¹⁶⁵³. Per exercir l'acció, llavors, només es necessita que l'entitat consideri que els drets de les persones interessades conferits en l'RGPD han estat vulnerats pel tractament de dades.

Podem veure aquí que el TJUE segueix la mateixa tendència tant en matèria antidiscriminatòria com en matèria de protecció de dades, és a dir, per un costat, no es necessita la concreció d'una víctima per poder dur a terme l'acció col·lectiva en matèria de protecció de dades (tal com en el cas Feryn, Accept i Rete Lenford en matèria antidiscriminatòria) i, per l'altre, l'acció no està supeditada a una vulneració concreta del dret a la protecció de dades sinó que és suficient l'existència d'un dany hipotètic (tal com en el cas Accept i Rete Lenford en matèria antidiscriminatòria).

En el cas de la discriminació algorítmica, en l'RGPD trobem diversos articles que habilitarien la presentació d'una demanda relativa a la discriminació algorítmica en una acció col·lectiva per vulneració del dret a la protecció de dades. Primer, en el cas d'existir una violació de l'article 22 de l'RGPD perquè s'ha dut a terme una decisió únicament automatitzada sense les garanties adequades i s'ha causat discriminació algorítmica. Segon, en cas que no hi hagués una decisió únicament automatitzada, es podria argumentar que els resultats generats per dades d'entrenament esbiaixades violen el principi d'adequació de les dades (art. 5.1.d de l'RGPD), o que la discriminació algorítmica no justificada també infringeix l'equitat en el tractament de dades (art. 5.1.a de l'RGPD)¹⁶⁵⁴.

En segon lloc, l'article 55 de la Directiva 2016/680 estableix que la persona interessada podrà donar mandat a una entitat que tingui per objecte protegir els drets i interessos de les persones interessades en relació amb les seves dades personals i que estigui constituïda d'acord amb la normativa de l'estat membre, per tal que presenti, en el seu nom, una reclamació davant l'autoritat de control i exerceixi el recurs judicial. De la mateixa manera que en l'RGPD, es podria argumentar, en el cas d'existir una violació de l'article 11 de la Directiva 2016/680, que s'ha dut a terme una decisió únicament automatitzada

¹⁶⁵³ TJUE (Sala Tercera), sentència de 28 d'abril de 2022, C-319/20, Meta Platforms Ireland Limited, anteriorment Facebook Ireland Limited, contra Bundesverband der Verbraucherzentralen und Verbraucherverbände-Verbraucherzentrale Bundesverband e.V., par. 83.

¹⁶⁵⁴ HACKER, P., "Teaching fairness to artificial intelligence...", *op. cit.*, 2018, pàg. 1168.

sense les garanties adequades i s'ha causat discriminació. En cas que no hi hagués una decisió únicament automatitzada, es podria argumentar que els resultats generats per dades d'entrenament esbiaixades violen el principi d'adequació de les dades (art. 4.1.c) o que la discriminació algorítmica no justificada també infringeix l'equitat en el tractament de dades (art. 4.1.a).

En tercer lloc, la Directiva 2020/1828 explícitament inclou, entre les disposicions del dret de la UE que si són infringides per part d'una persona empresària poden suposar l'exercici d'una acció de representació (dins l'àmbit del consum), les disposicions que es troben en l'RGPD¹⁶⁵⁵. Pel que fa a la discriminació algorítmica produïda arran d'una violació d'alguna disposició de l'RGPD en una relació de consum, la Directiva 2020/1828 possibilita que en tots els estats membres de la UE pugui litigar-se col·lectivament al respecte i desenvolupa quins són els continguts mínims que s'han d'assegurar en el mecanisme col·lectiu de protecció de les persones consumidores que s'estableixi en cada estat membre, com per exemple, concreta mesures de cessació¹⁶⁵⁶ i mesures reparadores¹⁶⁵⁷, estableix que la representació s'haurà de fer a través d'una entitat habilitada¹⁶⁵⁸, estableix requisits d'informació sobre les accions de representació en l'espai web de les entitats habilitades¹⁶⁵⁹, estableix que només s'ha de provar incompliment normatiu, sense necessitat de provar dol o negligència per part de l'empresari¹⁶⁶⁰, o possibilita que l'òrgan jurisdiccional o l'autoritat administrativa pugui ordenar a una part que exhibeixi certes proves demanades per l'altra part¹⁶⁶¹.

3.2. Acció de representació de protecció de dades a Espanya en el context de la discriminació algorítmica

A l'ordenament jurídic espanyol, com hem vist, es contempla l'acció col·lectiva en matèria de discriminació per interessos (pluri)individuals homogenis, col·lectius o

¹⁶⁵⁵ Apartat 56 de l'annex I de la Directiva 2020/1828.

¹⁶⁵⁶ Article 8 de la Directiva 2020/1828.

¹⁶⁵⁷ Article 9 de la Directiva 2020/1828.

¹⁶⁵⁸ Article 4 de la Directiva 2020/1828.

¹⁶⁵⁹ Article 13 de la Directiva 2020/1828.

¹⁶⁶⁰ Article 8.3 de la Directiva 2020/1828.

¹⁶⁶¹ Article 18 de la Directiva 2020/1828.

difusos, de manera que en qualsevol cas que ocorri una discriminació algorítmica les persones afectades podran emprar-la. Igualment, en el cas que la discriminació algorítmica es produeixi arran d'una infracció en matèria de protecció de dades, també es podrà recórrer a una acció col·lectiva en matèria de protecció de dades.

Quant a l'article 80 de l'RGPD, té efecte directe, de manera que si es produeix una infracció de l'RGPD que ocasioni discriminació algorítmica, una entitat, organització o associació amb mandat de la persona interessada podrà fer valer davant els tribunals de justícia espanyols els drets conferits a les persones individuals en l'RGPD¹⁶⁶². En el segon apartat de l'article 80, s'estableix que l'estat membre podrà disposar que una entitat, organització o associació sense mandat de la persona interessada també pugui exercir els drets conferits en el Reglament davant una autoritat de control o òrgan jurisdiccional.

Com ja hem apuntat en el l'apartat 5.1 del capítol V, l'article 80 de l'RGPD no ha estat objecte de desenvolupament en la LOPDGDD. L'LOPDGDD, únicament en el seu article 12.1, disposa que els drets reconeguts en l'article 15 a 22 de l'RGPD poden exercir-se directament o per mitjà de representant legal o voluntari¹⁶⁶³. Si ens adrecem llavors a la regulació general que trobem a LEC¹⁶⁶⁴, els mecanismes previstos per a la representació col·lectiva tenen per objecte la protecció de drets i interessos de les persones consumidores i no la protecció dels drets que dimanen de la LOPDGDD (essent dos àmbits diferents), de manera que la seva aplicació operaria en els casos en els quals un tractament il·lícit de dades personals podria també considerar-se una infracció que perjudiqués els interessos col·lectius de les persones consumidores¹⁶⁶⁵. A més, cal puntualitzar que l'RGPD exclou expressament que les entitats, organitzacions o associacions a les que fa referència a l'article 80 puguin reclamar una indemnització en nom de la persona interessada al marge d'aquesta última¹⁶⁶⁶, fet que no es dona en l'ordenament jurídic espanyol. Tal com exposa Villalba Cano, la desídia del poder legislatiu espanyol a l'hora de desenvolupar l'article 80 de l'RGPD en l'ordenament jurídic nacional ha donat lloc a un escenari jurídic complex i confús, no exempt

¹⁶⁶² Article 288 del TFUE. Aquest article disposa que el Reglament tindrà un abast general, serà obligatori en tots els seus elements i directament aplicable en cada estat membre.

¹⁶⁶³ COTINO HUESO, L., "Nuevo paradigma en la garantías...", *op. cit.*, 2022, pàg. 84.

¹⁶⁶⁴ FERNANDEZ SAMANIEGO, J. i PIÑAR GUZMÁN, B., "Las acciones colectivas en el marco del RGPD: una perspectiva desde el Derecho Civil español", *Diario La Ley*, núm. 26, Sección Ciberderecho, 11 de Febrero de 2019, pàg. 2.

¹⁶⁶⁵ *Idem, ibidem*.

¹⁶⁶⁶ VILLALBA CANO, L., "La representación de los interesados (Comentario al artículo 80 RGPD)", *op. cit.*, 2021, pàg. 3040.

d'incògnites i contradiccions derivades de la dificultat d'incloure allò disposat en la normativa europea de protecció de dades amb el règim preexistent que estableix la LEC, limitat a l'àmbit de la defensa dels interessos col·lectius de persones consumidores i usuàries¹⁶⁶⁷. Per tant, a Espanya, les entitats, organitzacions i associacions tindran la possibilitat de dur a terme una acció col·lectiva sense mandat de la persona interessada quan es tracti d'infraccions de protecció de dades que afectin a una relació de consum i per tant recaiguin en l'article 11 de la LEC.

Quant a l'article 55 de la Directiva 2016/680, el trobem transposat en l'article 55 de la Llei Orgànica 7/2021¹⁶⁶⁸, on es possibilita que la persona interessada confereixi la seva representació a una entitat per exercir el seu dret a recórrer a la jurisdicció contenciós-administrativa.

Quant a la Directiva 2020/1828, inclou explícitament que si les disposicions que es troben en l'RGPD són infringides per part d'una persona empresària poden suposar l'exercici d'una acció de representació (dins l'àmbit del consum)¹⁶⁶⁹. De manera que una infracció de l'RGPD en l'àmbit del consum, a través de la Directiva 2020/1828, també dona possibilitat a l'acció col·lectiva prevista en l'article 11 de la LEC¹⁶⁷⁰.

Per concloure, l'acció col·lectiva en matèria de protecció de dades, a Espanya, no proporciona cap millora en relació amb l'acció col·lectiva en matèria de discriminació

¹⁶⁶⁷ *Ídem*, pàg. 3034.

¹⁶⁶⁸ Article 55 de la Llei Orgànica 7/2021.

¹⁶⁶⁹ Apartat 56 de l'annex I de la Directiva 2020/1828.

¹⁶⁷⁰ La legitimació per les accions col·lectives en defensa d'interessos de persones consumidores (art. 11 de la LEC), difereix de la legitimació en defensa d'interessos de persones que s'han vist discriminades (art. 11 bis de la LEC). S'estableixen en el primer cas com a legitimades (a part del Ministeri Fiscal que està legitimat per exercitar qualsevol acció en defensa de persones consumidores i usuàries (art. 11.5 LEC)) les associacions de persones consumidores i usuàries legalment constituïdes per defensar tant a les persones associades com els interessos generals de les persones consumidores i usuàries (art. 11.1 LEC). Si el grup està perfectament determinat o és fàcilment determinable, la legitimació per la defensa dels interessos col·lectius correspon a les associacions de persones consumidores i usuàries, a les entitats legalment constituïdes que tinguin per objecte la defensa i protecció d'aquests i els propis grups afectats (art. 11.2 LEC). Si les persones perjudicades per un fet perjudicial són una pluralitat indeterminada o de difícil determinació, la legitimació correspondrà exclusivament a les associacions de persones consumidores i usuàries que conforme a la llei siguin representatives (art. 11.3 LEC) i a les entitats habilitades segons la normativa europea per les accions de cessació (art. 11.4 LEC). En canvi, la legitimació per dur a terme accions col·lectives en defensa d'interessos difusos en matèria antidiscriminació, segons l'article 11 bis.2 de la LEC legitima a l'Autoritat Independent per a la Igualtat de Tracte i la No Discriminació, partits polítics, sindicats, associacions professionals de treballadors autònoms més representatives, organitzacions de persones consumidores i usuàries d'àmbit estatal, organitzacions, d'àmbit estatal o de l'àmbit territorial on es produeixi la situació de discriminació que tinguin entre les seves finalitats la defensa i promoció dels drets humans, sense perjudici en tot cas de la legitimació individual d'aquelles persones afectades que estiguessin determinades.

per defensar interessos de persones que s'han vist discriminades mitjançant sistemes algorítmics.

En canvi, com veurem seguidament, en relació amb l'obtenció de prova, l'RGPD i la Directiva 2020/1828 sí que proporcionen més eines per obtenir prova que la normativa antidiscriminació, i poden ser rellevants en accions col·lectives per cessar conductes que generin discriminació algorítmica i obtenir reparació.

3.3. Aportació de la normativa europea de protecció de dades a l'acció de representació per la lluita envers la discriminació algorítmica

L'RGPD, a diferència de la normativa antidiscriminatòria, proporciona certs drets individuals que com hem explicat poden conduir a obtenir prova que s'ha produït discriminació algorítmica. En el cas d'haver-se produït la discriminació a través d'una decisió automatitzada, els articles 13.2 (f), 14.2 (g) i 15.1 (h) estableixen que la persona responsable del tractament facilitarà a la persona interessada informació sobre l'existència d'una decisió automatitzada o l'elaboració de perfils, informació significativa sobre la lògica aplicada, i la importància i les conseqüències previstes del tractament per a la persona interessada. Tal com hem esmentat en l'apartat 3.2 del capítol V, creiem que per tal que aquest tipus d'explicació pugui ser eficient en casos de discriminació algorítmica, s'hauria d'articular en tant que explicació contrafàctica.

D'aquesta manera, aquests drets que atorga l'RGPD (individualment), si es presentés la informació obtinguda per cada persona de forma col·lectiva, podrien demostrar un patró o tendència del sistema algorítmic a discriminar cert grup determinat de població, facilitant (pel fet de presentar-ho conjuntament -aquest seria el cas d'interessos (pluri)individuals homogenis-, o pel fet de plantejar que s'ha donat el cas en diverses persones pertanyents a un col·lectiu determinat -aquest seria el cas d'interessos col·lectius, difusos o l'interès general-) l'establiment d'indicis de discriminació *prima facie*, generant, conseqüentment, la inversió de la carrega de la prova.

Per demostrar aquest patró o tendència del sistema algorítmic a discriminar certs grups de població, en aquest treball (vegeu l'apartat 4.3.3 del capítol V) també ens hem referit,

en base a l'RGPD, al requeriment d'avaluacions d'impacte algorítmiques, és a dir, avaluacions que anessin més enllà de les avaluacions d'impacte de protecció de dades. Aquestes avaluacions d'impacte algorítmiques descriurien, per exemple, poblacions o situacions en les quals és menys probable que el sistema algorítmic funcioni de manera òptima¹⁶⁷¹ i podrien ajudar a exercir drets col·lectius. No obstant això, haurien de ser públiques i accessibles. Com tractarem en el següent apartat, sembla que sí que es podrà obtenir informació de rellevància sobre el funcionament del sistema algorítmic en base a allò estipulat en la PRIA, tot i que no serà accessible al públic general.

També creiem rellevant destacar l'article 18 de la Directiva 2020/1828 que, atès que és evident que en l'àmbit del consum hi ha un clar desequilibri de poder entre empresa i persona usuària (com també en el cas de la discriminació algorítmica entre empresa o organització i persona perjudicada), estableix que els estats membres vetllaran per tal que quan una entitat habilitada¹⁶⁷² hagi aportat proves raonablement disponibles, suficients per donar suport a una acció de representació, i hagi indicat altres proves que estiguin en poder de la part demandada o d'una tercera part, si així ho demana l'entitat habilitada, l'òrgan jurisdiccional o l'autoritat administrativa pugui ordenar a la part demandada o tercera part que exhibeixi les proves sol·licitades. També podrà la part demandada demanar l'exhibició de proves en possessió de la part demandant o d'una tercera part¹⁶⁷³.

Segons el considerant 68, aquest article respon a que les relacions entre empresa i persona consumidora es caracteritzen habitualment per l'asimetria d'informació o al fet que les proves necessàries poden estar exclusivament en poder de la persona empresària, de manera que són inaccessibles per a l'entitat habilitada, i per tant, aquesta última ha de tenir dret a demanar a la persona empresària o a l'autoritat administrativa que ordeni a la persona empresària que exhibeixi les proves pertinents per la seva pretensió. Pel principi d'igualtat d'armes, la persona empresària també ha de tenir un dret similar. L'òrgan jurisdiccional o l'autoritat administrativa ha de valorar la necessitat, abast i

¹⁶⁷¹ SELBST, A. D., "An institutional view of Algorithmic Impact Assessments", *op. cit.*, 2021, pàg. 137; KAMINSKI, M. E. i MALGIERI, G., "Algorithmic impact assessments...", *op. cit.*, 2021, pàg. 138.

¹⁶⁷² La Directiva 2020/1828 defineix entitat habilitada, segons l'article 3.4, com tota organització o organisme públic que representi els interessos de les persones consumidores, que hagi estat designada com a tal per un estat membre per l'exercici d'accions de representació de conformitat amb la directiva.

¹⁶⁷³ Article 18 de la Directiva 2020/1828. La transposició d'aquest article a Espanya, segons l'Anteproyecto de Ley de acciones de representación para la protección de los intereses colectivos de los consumidores, es farà amb la introducció de l'article 838 a la LEC.

proporcionalitat de l'exhibició de proves¹⁶⁷⁴.

En el cas de la discriminació algorítmica, tal com comentàvem en relació amb l'article 4.b de la Propuesta de Directiva del Consejo relativa a la carga de la prueba en el ámbito de la igualdad de retribución y trato entre hombres y mujeres¹⁶⁷⁵, caldria un article que possibilités l'accés a informació determinant en relació amb la precisió del sistema en diferents col·lectius determinats de població que deté l'empresa que ha creat un sistema algorítmic, ja que es tracta d'un element essencial per provar que s'ha ocasionat una discriminació. Tanmateix, actualment, aquest article només seria aplicable en els casos en els quals la discriminació algorítmica s'hagués produït a causa d'una infracció que es trobés dins l'àmbit d'aplicació de la Directiva 2020/1828.

Recapitulant, la normativa i jurisprudència europea també possibiliten una acció col·lectiva en matèria de protecció de dades en el cas que un tractament de dades il·lícit provoqui discriminació algorítmica. No obstant això, a Espanya no s'ha profunditzat en l'acció col·lectiva de protecció de dades i només està habilitada en el cas d'una relació de consum. En relació amb la prova de la discriminació algorítmica, si la discriminació es produeix per la il·licitud d'un tractament de dades en una relació de consum, i es troba en l'àmbit de l'acció de representació de la Directiva 2020/1828, l'article 18 obliga els estats membres a que quan l'entitat habilitada hagi aportat proves raonablement disponibles, i hagi indicat altres proves que estiguin en el poder de la part demandada o d'una tercera part, l'òrgan jurisdiccional o l'autoritat administrativa pugui ordenar a la part demandada o tercera part que exhibeixi les proves sol·licitades, de manera que la part demandant podria obtenir informació sobre la precisió del sistema algorítmic en certs grups determinats de població o qualsevol informació pertinent per provar que s'ha produït la discriminació algorítmica. Tot seguit analitzem què pot aportar la PRIA a l'acció col·lectiva en matèria de discriminació.

¹⁶⁷⁴ Considerant 68 de la Directiva 2020/1828. Tal com en l'e-Discovery americà, cada una de les parts podria sol·licitar al tribunal que es requerís a l'altra part la preservació d'evidències electròniques que estan en el seu poder, per tal de poder-les revisar posteriorment i utilitzar com a prova.

¹⁶⁷⁵ COMISIÓN DE LAS COMUNIDADES EUROPEAS, *Propuesta de Directiva del Consejo relativa a la carga de la prueba en el ámbito de la igualdad...*, op. cit., 27 de mayo de 1988, art. 4.b.

4. PRIA i acció de representació en matèria de discriminació algorítmica

La PRIA no preveu mecanismes de representació col·lectiva ni tampoc atorga drets individuals (malgrat que això pot canviar en el text final arran de les esmenes proposades pel Parlament Europeu¹⁶⁷⁶). Tanmateix, obliga els proveïdors de sistemes de risc alt a crear documentació tècnica¹⁶⁷⁷ i habilita els organismes nacionals d'igualtat a accedir a certa informació a la qual no dona accés a persones individualment considerades¹⁶⁷⁸.

La documentació tècnica, com ja hem esmentat, requereix aportar informació sobre la precisió del sistema sobre certs grups de població. Interessa puntualitzar que només es requereix la presentació d'aquesta documentació per als sistemes de risc alt, per tant, aquells que no ho siguin quedaran exclosos de la possibilitat que un organisme d'igualtat nacional pugui emprar la informació que allà es proporciona. L'annex IV de la PRIA, en relació amb l'article 11, com ja hem especificat en l'apartat 3.2 del capítol VI del treball, proporciona la informació mínima que haurà de constar en la documentació tècnica¹⁶⁷⁹. Entre aquesta informació es troba, per exemple, el requisit de proporcionar informació detallada sobre les limitacions de funcionament, inclosos els nivells de precisió per a les persones o grups de persones específics respecte als quals es preveu emprar el sistema i el nivell de precisió general esperat (apartat 3 de l'annex IV), informació que pot ser molt rellevant en el cas que s'hagin de demostrar indicis de discriminació *prima facie* per part d'un sistema algorítmic.

Sota el nostre parer, els sistemes algorítmics que en aquesta documentació tècnica mostressin menys precisió segons el grup determinat de població al qual s'apliquessin haurien d'estar prohibits, atesa la importància de les repercussions que s'ha demostrat que pot tenir un sistema algorítmic que discrimini. En canvi, no només no estan prohibits sinó que la documentació que pot servir per provar que el sistema algorítmic, pel seu funcionament, és menys precís en certs grups de població, no estarà disponible per ser consultada per part de les persones que es veuran discriminades o intuiran que s'han vist discriminades, sinó que l'articulat de la PRIA estableix només l'accés per les autoritats

¹⁶⁷⁶ EUROPEAN PARLIAMENT, *Amendments adopted by the European Parliament on 14 June 2023...*, *op. cit.*, 14 June 2023.

¹⁶⁷⁷ Article 11 de la PRIA.

¹⁶⁷⁸ Article 64.3 i considerant 79 de la PRIA.

¹⁶⁷⁹ Article 11 de la PRIA.

nacionals competents i els organismes notificats, i únicament per avaluar si el sistema de risc alt compleix amb els requisits del Reglament. L'article 64.3 i considerant 79 de la PRIA, tanmateix, estipulen que les autoritats o organismes públics nacionals que supervisin l'aplicació del dret de la UE que protegeix els drets fonamentals, inclosos els organismes d'igualtat, també tindran accés a la documentació que es creï en virtut de la PRIA¹⁶⁸⁰. Cotino Hueso planteja que quan les autoritats disposin de la documentació tècnica que la PRIA obliga a confeccionar, en molts casos -sobretot quan es tracti d'usos públics de la IA- es tractarà d'informació pública¹⁶⁸¹ a la qual es podrà accedir a través del dret d'accés a informació pública¹⁶⁸². Sembla aquesta una bona proposta per tal que les persones puguin accedir a la documentació tècnica dels sistemes d'IA.

També, l'article 60 de la PRIA estableix que la Comissió, en col·laboració amb els estats membres, crearà i mantindrà una base de dades de la UE dels sistemes de risc alt que encaixin amb els paràmetres establerts en l'annex III de la PRIA¹⁶⁸³. A través d'aquesta base de dades pública, les persones podran localitzar si el sistema algorítmic a través del qual han estat subjectes d'una decisió s'integra en un sistema de risc alt que hagi generat documentació tècnica. Igualment, per tal que aquest sistema funcioni, la persona ha de saber exactament en base a quin sistema algorítmic s'ha pres la decisió de la qual ha estat subjecte, per tant, les empreses i administracions públiques que n'emprin haurien de publicar sempre que prenguin un decisió mitjançant un sistema algorítmic, en base a quin sistema l'han pres, ja sigui propi o extern i que aquest estigui registrat. Prèviament, seria rellevant que tots els sistemes algorítmics estiguessin identificats.

A tall de recopilació, la PRIA no dona drets col·lectius ni estableix mecanismes d'acció col·lectiva envers sistemes algorítmics discriminatoris, de manera que si això no canvia en el text final i s'atorguen drets a les persones envers els sistemes algorítmics, únicament podria fer aportacions quant a la prova de discriminació algorítmica. Tant pot ser útil l'accés a la documentació tècnica per part dels organismes d'igualtat, com l'establiment d'una base de dades pública que faciliti la identificació de sistemes algorítmics. Anant

¹⁶⁸⁰ Article 64.3 i considerant 79 de la PRIA.

¹⁶⁸¹ L'article 13 de la Llei 19/2013 disposa que: *Se entiende por información pública los contenidos o documentos, cualquiera que sea su formato o soporte, que obren en poder de alguno de los sujetos incluidos en el ámbito de aplicación de este título y que hayan sido elaborados o adquiridos en el ejercicio de sus funciones.*

¹⁶⁸² COTINO HUESO, L. "Qué concreta transparencia e información de algoritmos e inteligencia artificial es la debida", *Revista Española de la Transparencia (RET)*, núm. 16, 2023, pàg. 57.

¹⁶⁸³ Article 60.1 de la PRIA.

més enllà, seria pertinent que una norma exigís a qualsevol empresa o organització pública que emprés sistemes algorítmics per a la presa de decisions amb impacte en les persones, que publiqués que n'està emprant i amb quin objectiu.

5. Conclusions parcials

En aquest capítol ens proposàvem respondre, en part, a la cinquena qüestió de recerca del treball, concretament, si és necessària la promoció de la representació col·lectiva per fer front a la discriminació algorítmica.

L'acció col·lectiva és adequada per adreçar la discriminació algorítmica per diverses raons. És un instrument que possibilita la protecció dels interessos col·lectius afectats (no prioritaris en la normativa antidiscriminació europea basada clarament en la protecció individual), atès que tot i que les conseqüències d'una acció discriminatòria puguin ser individuals, són els estereotips i prejudicis contra el col·lectiu al qual pertany una persona els que ocasionen el resultat discriminatori.

L'ús de l'acció col·lectiva posa en rellevància el component grupal de la discriminació, i millora les possibilitats d'establir un cas *prima facie*, en haver més persones legitimades per desafiar la discriminació algorítmica. A més, si múltiples persones que pertanyen a un col·lectiu de població determinat que comparteix una (o diverses) categoria(es) protegida(es) plantegen que s'han vist perjudicades per un mateix sistema algorítmic, serà més fàcil evidenciar que el tracte rebut per part del sistema algorítmic és sistemàticament pitjor a causa de la pertinença a la/les categoria/es protegida/es. També millora les possibilitats d'establir *tertium comparationis*, ja que hi ha entitats legitimades que tenen més possibilitats d'accedir a informació sobre el sistema algorítmic que no pas les persones individualment considerades.

L'acció col·lectiva en matèria de discriminació algorítmica podria dur-se a terme per protegir interessos col·lectius, difusos, (pluri)individuals homogenis o l'interès general i podria tenir l'objectiu tant de cessar l'acció discriminatòria com de reparar els perjudicis causats. A la UE no hi ha un mecanisme harmonitzat d'acció col·lectiva antidiscriminació. Únicament les directives antidiscriminació obliguen els estats membres a que

associacions, organitzacions i entitats jurídiques puguin iniciar en nom de la persona demandant o en el seu suport i autorització procediments judicials i administratius. Els estats membres poden oferir més possibilitats, com és el cas d'Espanya, on és possible la defensa d'interessos difusos de persones que s'han vist discriminades, no essent únicament la persona discriminada que ha de vetllar per la seva protecció, sinó que entitats amb més accés a informació i recursos, per exemple els organismes d'igualtat nacionals, poden desafiar decisions que causin discriminació algorítmica i plantejar la cessació de l'algoritme que causa la discriminació.

La normativa i jurisprudència en matèria de protecció de dades també possibiliten una acció col·lectiva quan un tractament de dades il·lícit provoqui discriminació algorítmica. No obstant això, a Espanya no s'ha aprofundit en l'acció col·lectiva de protecció de dades i només està habilitada en el cas que es doni dins d'una relació de consum. En aquests casos, l'article 18 de la Directiva 2020/1828 relativa a les accions de representació en matèria de dret de les persones consumidores obliga els estats membres a que quan l'entitat habilitada hagi aportat proves raonablement disponibles, i hagi indicat altres proves que estiguin en el poder de la part demandada o d'una tercera part, l'òrgan jurisdiccional o l'autoritat administrativa pugui ordenar a la part demandada o tercera part, que exhibeixi les proves sol·licitades. Aquest article obre la possibilitat a la part demandant per obtenir informació sobre la precisió del sistema algorítmic en certs grups determinats de població o qualsevol informació pertinent per provar que s'ha produït la discriminació algorítmica.

La PRIA, en canvi, no atorga drets col·lectius ni estableix mecanismes d'acció col·lectiva envers sistemes algorítmics discriminatoris (malgrat que això pot canviar en el text final arran de les emenes proposades pel Parlament Europeu¹⁶⁸⁴). No obstant això, facilita l'accés a la documentació tècnica dels sistemes d'IA de risc alt per part dels organismes d'igualtat -possibilitant que a través d'aquesta informació es pugui evidenciar que la precisió del sistema algorítmic és comparativament millor o pitjor segons el grup determinat de població-, i estableix una base de dades pública que facilita la identificació de sistemes algorítmics. És a dir, ofereix informació rellevant per la defensa de drets col·lectius.

¹⁶⁸⁴ EUROPEAN PARLIAMENT, *Amendments adopted by the European Parliament on 14 June 2023...*, *op. cit.*, 14 June 2023.

Per tant, concloem que és necessària la promoció de la representació col·lectiva per fer front a la discriminació algorítmica. No obstant això, és un instrument que plantejat per si sol presenta certes limitacions. Detectem que, en l'estat actual de la qüestió, en el qual les persones tenen molt poc accés a informació de rellevància, estan molt disseminades i isolades, i són poc conscients o estan poc preocupades pels riscos que planen sobre la seva vida privada, sense conèixer les noves capacitats de vigilància que té la intel·ligència artificial, tant respecte la vida laboral com a la vida social i privada¹⁶⁸⁵, s'ha d'estudiar com els organismes nacionals d'igualtat (que tenen facultats d'investigació i en la PRIA se'ls dona la possibilitat d'obtenir certa informació sobre sistemes algorítmics) han d'integrar en les seves funcions l'actuació en benefici de les persones afectades per discriminació algorítmica i inclús, en molts casos, actuar d'ofici. Considerem llavors que és necessari incentivar la institucionalització de l'acció col·lectiva, atès que no es podrà brindar una protecció efectiva sense un organisme públic que pugui, en primer lloc, accedir a la informació sobre l'impacte precís del sistema algorítmic sobre grups determinats de persones; en segon lloc, centralitzar (telemàticament) les demandes de les persones afectades per un sistema algorítmic discriminatori i gestionar i promoure una acció de cessació del sistema algorítmic discriminatòria (a través de la combinació entre la informació individual i la informació sobre la precisió del sistema algorítmic en determinats grups de població); i en tercer lloc, posar en contacte comunitats afectades i proveïdors d'aquests sistemes per col·laborar en crear sistemes més sensibles a les conseqüències discriminatòries. Aquestes seran les qüestions que abordem en el següent capítol.

¹⁶⁸⁵TORRALBA, F., *L'ètica algorítmica, op. cit.*, 2022, pàg. 170.

Capítol VIII. Protecció institucional: organismes nacionals d'igualtat com a actors necessaris per fer front a la discriminació algorítmica

Després d'estudiar les possibilitats de l'acció col·lectiva en matèria de discriminació algorítmica per a la protecció dels interessos de persones i col·lectius discriminats mitjançant sistemes algorítmics, en aquest capítol, per respondre a la segona part de la cinquena qüestió de recerca, a saber, si cal una institució pública que coordini la lluita envers la discriminació algorítmica, estudiem la necessitat d'una protecció institucional que complementi la protecció col·lectiva per crear un sistema de protecció més efectiu.

Aquesta protecció institucional pot assegurar l'operativitat de la representació col·lectiva interrelacionant persones discriminades per sistemes algorítmics i adoptant un rol proactiu -per exemple, actuant sense víctimes identificades- per cessar sistemes algorítmics discriminatoris. Defensem que s'ha de relacionar indefectiblement la defensa de les persones discriminades amb el sector públic (que vetlla per l'interès general i no per l'interès particular), i que aquesta relació s'ha de traduir en un suport institucional proactiu. Per tant, en aquest capítol descriuim una de les eines que poden facilitar el desenvolupament de noves estratègies reguladores efectives en l'àmbit de la discriminació algorítmica¹⁶⁸⁶, estant aquesta eina directament relacionada amb oferir una oportunitat de defensa a les persones i col·lectius que s'han vist subjectes a discriminació mitjançant sistemes algorítmics.

Per consegüent, estructurarem el capítol amb un primer apartat on plantejarem que són els organismes d'igualtat que els estats membres de la UE estan obligats a tenir els més ben posicionats per oferir una millor protecció envers la discriminació algorítmica (1), concretant en el segon apartat quines funcions han de dur a terme per crear una protecció efectiva de les persones que s'hagin vist discriminades (2), especificant, en el tercer apartat, si l'organisme d'igualtat espanyol constituït per la Llei 15/2022, és a dir, l'Autoritat Independent, s'ha instituït d'acord amb les directrius europees i pot dur a terme un suport institucional efectiu envers la discriminació algorítmica (3), i abordant, en el

¹⁶⁸⁶ Vegeu, en relació amb l'article que inspira la proposta que formulem en aquest capítol: KOULU, R., "Human Control over Automation: EU Policy and AI Ethics", *European Journal of Legal Studies*, vol. 12, núm. 1, 2020, pàg. 44. L'autora es pregunta què podem i què hem de regular i com poden els estudis sociojurídics facilitar el desenvolupament de noves estratègies reguladores efectives.

quart apartat, la complementarietat i la cooperació entre l'Agència Espanyola de Supervisió de la Intel·ligència Artificial (d'ara en endavant, AESIA), l'AEPD i l'Autoritat Independent, atesa la implicació dels dos primers organismes en el control *ex ante* i *ex post* dels sistemes d'IA que empren algoritmes que poden produir resultats discriminatoris (4). Finalment, acabarem el capítol amb un apartat de conclusions parcials (5).

1. Els organismes d'igualtat dels estats membres de la UE com a òrgans més ben posicionats per desafiar decisions discriminatòries preses mitjançant sistemes algorítmics

Els organismes d'igualtat de la UE, sota la normativa europea, tenen un rol fonamental en la supervisió de l'aplicació de la normativa antidiscriminació i, per les competències que tenen atorgades, poden jugar un paper essencial en el desafiament de decisions preses mitjançant sistemes algorítmics¹⁶⁸⁷. En aquest sentit, per exemple, poden assumir o alleugerir la responsabilitat de detectar la discriminació i fer-li front que actualment es diposita en la persona discriminada per un sistema algorítmic¹⁶⁸⁸, poden ser multifuncionals, centralitzar la informació de moltes persones diferents i, atès que el seu objectiu és actuar en benefici de l'interès general, poden assegurar que no seran instrumentalitzats per altres finalitats que no siguin la defensa de la no discriminació (és a dir, evitar la possibilitat que existeixi, per exemple, un conflicte d'interès comercial¹⁶⁸⁹), sempre que s'estableixin bones salvaguardes per la seva independència¹⁶⁹⁰ -tant funcional com econòmica. Tanmateix hi ha certs factors que determinen la seva efectivitat i com apunten Chopin i Germaine, malgrat que la igualtat formal s'ha obtingut en la majoria de legislacions nacional, s'han de fer més esforços per aconseguir la igualtat substantiva¹⁶⁹¹.

Per poder dur a terme la seva tasca en relació amb els sistemes d'intel·ligència artificial,

¹⁶⁸⁷ Per una defensa de la rellevància dels organismes d'igualtat nacionals (independents i amb recursos) com a eina fonamental que pot ser usada per adreçar la discriminació, vegeu: BENEDI LAHUERTA, S., "Equality bodies: advancing towards more responsive designs?", *International Journal of law in context*, vol. 17, núm. 3, 2021, pàg. 407.

¹⁶⁸⁸ HAKKARAINEN, J., "Naming something collective...", *op. cit.*, 2021, pàg. 18 i 19.

¹⁶⁸⁹ HODGES, C., "Collective Redress: The Need for New Technologies", *op. cit.*, 2019, pàg. 87 i 88.

¹⁶⁹⁰ Sobre la independència de les agències administratives, vegeu, per exemple: SCHERER, M. U., "Regulating Artificial Intelligence Systems: Risks, Challenges, Competencies, and Strategies", *Harvard Journal of Law & Technology*, vol. 29, núm. 2, 2016, pàg. 385 i 386.

¹⁶⁹¹ CHOPIN, I. i GERMAINE, C., A comparative analysis of..., *op. cit.*, 2022, pàg. 131.

com apunten Allen i Masters, necessiten guia sobre com aplicar la legislació existent en matèria d'igualtat en situacions que involucren l'ús de tecnologies de la intel·ligència artificial¹⁶⁹² i han de de revisar els seus mandats per assegurar que tenen poders significatius i adequats per adreçar els nous reptes de la IA i el seu desafiament al principi de no-discriminació¹⁶⁹³.

En aquest apartat analitzem la configuració dels organismes d'igualtat dels estats membres de la UE, d'una primera fase on comptaven amb un àmbit d'actuació molt limitat (1.1) a una segona fase on s'està produint una expansió tant del seu àmbit d'aplicació, com de les seves funcions i competències (1.2)

1.1. Primera fase: àmbit d'actuació reduït i poques competències

La creació d'organismes d'igualtat en els estats membres ve determinada per l'article 13 de la Directiva 2000/43/CE que disposa, en el seu primer apartat, que cada estat membre designarà un o més organismes responsables de la promoció de la igualtat de tracte entre totes les persones sense discriminació per raó d'origen ètnic¹⁶⁹⁴. De la mateixa manera, posteriorment, l'article 12.1 de la Directiva 2004/113/CE¹⁶⁹⁵, l'article 20.1 de la Directiva 2006/54/CE¹⁶⁹⁶ i l'11.1 de la Directiva 2010/41/UE del Parlamento Europeu y del Consejo, de 7 de julio de 2010, sobre la aplicación del principio de igualdad de trato entre hombres y mujeres que ejercen una actividad autónoma (d'ara en endavant, 2010/41/UE¹⁶⁹⁷), estableixen articles en el mateix sentit en relació amb la creació d'un organisme responsable de la promoció de la igualtat de tracte entre totes les persones

¹⁶⁹² ALLEN, R. i MASTERS, D., *Regulating for an equal AI...*, op. cit., 2020, pàg. 11.

¹⁶⁹³ *Ídem*, pàg. 22.

¹⁶⁹⁴ Article 13.1 de la Directiva 2000/43/CE.

¹⁶⁹⁵ L'article 12.1 de la Directiva 2004/113/CE disposa que: "*Los Estados miembros designarán uno o más organismos para que se encarguen de promover, analizar, controlar y apoyar la igualdad entre todas las personas sin discriminación por razones de sexo, y adoptarán las disposiciones necesarias a tal fin*".

¹⁶⁹⁶ L'article 20.1 de la Directiva 2006/54/CE disposa que: "*Cada Estado miembro designará uno o más organismos responsables de la promoción, el análisis, el seguimiento y el apoyo de la igualdad de trato entre todas las personas, sin discriminación por razón de sexo, y adoptará en este sentido las disposiciones necesarias*".

¹⁶⁹⁷ L'article 11.1 de la Directiva 2010/41/UE disposa que: "*Cada Estado miembro adoptará las medidas necesarias para velar por que el organismo u organismos designados de conformidad con el artículo 20 de la Directiva 2006/54/CE sean asimismo competentes para la promoción, el análisis, el seguimiento y el apoyo de la igualdad de trato entre todas las personas incluidas en el ámbito de aplicación de la presente Directiva, sin discriminación por razón de sexo*".

sense discriminació per raó de sexe. En les directives s'empra el terme "designarà", de manera que tots els estats membres, en vista de l'obligatorietat de la directiva, han establert organismes d'igualtat¹⁶⁹⁸.

S'entén que l'organisme d'igualtat actuarà en els àmbits establerts per les directives, és a dir, en àmbits diferents segons la categoria protegida. En base a l'origen ètnic o racial, es requereix l'organisme d'igualtat en els àmbits del treball, inclòs el treball autònom, l'ocupació, l'educació, la protecció social (inclosa la seguretat social i l'assistència sanitària), avantatges socials, i accés i subministrament de béns i serveis (inclòs l'habitatge); en base al sexe, es requereix l'organisme d'igualtat en els àmbits del treball, inclòs el treball autònom, l'ocupació i accés i subministrament de béns i serveis¹⁶⁹⁹.

No hi ha cap article en aquest sentit en la Directiva 2000/78/CE¹⁷⁰⁰, i per tant, no existeix actualment una normativa a escala europea que obligui als estats membres a tenir un organisme d'igualtat responsable de la promoció de la igualtat de tracte entre totes les persones sense discriminació per raó de seva religió o conviccions, discapacitat, edat o orientació sexual. Tanmateix, segons la Comissió Europea, addicionalment a les seves obligacions en les directives esmentades, la majoria dels estats membres han ampliat el mandat dels seus organismes d'igualtat per abastar la discriminació per motius de sexe, origen racial o ètnic, convicció o religió, discapacitat, edat o orientació sexual en els àmbits del treball i l'ocupació, l'accés a béns i serveis i el seu subministrament, l'educació, la protecció social i els beneficis socials, entre altres, cobrint així l'aplicació de la Directiva 2000/78/CE¹⁷⁰¹.

Quant a les competències dels organismes d'igualtat, les quatre directives estableixen que aquests han de prestar assistència independent a les víctimes de discriminació iniciant un

¹⁶⁹⁸ Considerant 11 de la Recomendación (UE) 2018/951.

¹⁶⁹⁹ EUROPEAN COMMISSION, *Equality bodies and the implementation of the Commission Recommendation on standards for equality bodies Accompanying the document Report from the Commission to the European Parliament and Council on the application of Council Directive 2000/43/EC implementing the principle of equal treatment between persons irrespective of racial or ethnic origin ('the Racial Equality Directive') and of Council Directive 2000/78/EC establishing a general framework for equal treatment in employment and occupation ('the Employment Equality Directive')*, COM(2021) 139 final, Brussels, 19 May 2021, pàg. 4. Al llarg del treball anomenarem aquest document de la següent manera: Informe de la Comissió que avalua l'aplicació de la Recomanació 2018/951, de 22 de juny de 2018, de 19 de març de 2021.

¹⁷⁰⁰ Tanmateix, s'inclou una disposició obligant a la designació d'un organisme d'igualtat en l'article 12.1 de la Propuesta de Directiva del Consejo por la que se aplica el principio de igualdad de trato entre las personas independientemente de su religión o convicciones, discapacidad, edad u orientación sexual, de 2 de juliol de 2008.

¹⁷⁰¹ Considerant 17 de la Recomendación (UE) 2018/951.

procediment per discriminació, realitzant estudis independents sobre la discriminació i publicant informes independents, i formulant recomanacions sobre qualsevol qüestió relacionada amb la discriminació¹⁷⁰². Les directives 2006/54/CE i 2010/41/UE afegeixen que els organismes d'igualtat hauran d'intercanviar la informació disponible amb els organismes europeus equivalents¹⁷⁰³.

Molts estats membres han anat més enllà d'allò exigint en les directives, ja sigui en termes de les característiques protegides¹⁷⁰⁴, ja sigui en termes del poders atorgats per lluitar envers la discriminació¹⁷⁰⁵. Majoritàriament s'han instaurat organismes que protegeixen diverses característiques protegides alhora¹⁷⁰⁶. Les directives esmentades han deixat als estats membres amb un ampli marge d'apreciació, i per tant hi ha diferències significatives en relació amb el mandat, les competències, estructures, recursos i funcionament operacional, fet que ocasiona que la protecció sigui desigual segons l'estat membre¹⁷⁰⁷. A més, s'han identificat casos d'organismes d'igualtat que manquen d'independència i eficàcia¹⁷⁰⁸ per motius com l'existència de pressions exteriors¹⁷⁰⁹,

¹⁷⁰² Article 13.2 de la Directiva 2000/43/CE; article 12.2 de la Directiva 2004/113/CE; article 20.2 de la Directiva 2006/54/CE; i article 11.2 de la Directiva 2010/41/UE.

¹⁷⁰³ Article 20.2 de la Directiva 2006/54/CE i article 11.2 de la Directiva 2010/41/UE. Per una anàlisi comparativa sobre les competències dels organismes d'igualtat en tots els estats membres de la UE, específicament les competències de proveir assistència a les víctimes, dur a terme enquestes, publicar informes i fer recomanacions, dur a terme funcions quasi-jurisdiccionals i fer decisions vinculants, vegeu: CHOPIN, I. i GERMAINE, C., *A comparative analysis of...*, *op. cit.*, 2022, pàg. 119-121.

¹⁷⁰⁴ ALLEN, R. i MASTERS, D., *Regulating for an equal AI...*, *op. cit.*, 2020, pàg. 39. En aquest article s'afirma que molts estats membres tenen un conjunt coherent de drets per totes les categories protegides tant en l'àmbit de l'ocupació com en l'accés i subministrament a béns i serveis.

¹⁷⁰⁵ CROWLEY, N., *Equality bodies making a difference*, European network of legal experts in gender equality and non-discrimination, Brussels, 2018., pàg. 9. Per informació detallada sobre quins atributs protegeix i quines competències té cada organisme d'igualtat en els estats membres europeus, vegeu: CHOPIN, I. i GERMAINE, C., *A comparative analysis of...*, *op. cit.*, 2022, pàg. 109-118. S'ha de tenir en compte en relació amb aquesta informació, que en el cas d'Espanya encara es fa referència al CEDRE i no a l'Autoritat Independent creada per la Llei 15/2022.

¹⁷⁰⁶ *Ídem*, pàg. 9.

¹⁷⁰⁷ Considerant 18 de la Recomendación (UE) 2018/951.

¹⁷⁰⁸ CROWLEY, N., *Equality bodies making a difference*, *op. cit.*, 2018, pàg. 10-12. Per evidències sobre la manca d'independència i eficàcia (atesa la manca de recursos) dels organismes d'igualtat dels estats membres, vegeu: CHOPIN, I. i GERMAINE, C., *A comparative analysis of...*, *op. cit.*, 2022, pàg. 122-124.

¹⁷⁰⁹ CROWLEY, N., *Equality bodies making a difference*, *op. cit.*, 2018, pàg. 7. En aquest informe s'analitzen els 43 organismes d'igualtat europeus, i es conclou que el context polític per a vuit d'ells és d'hostilitat política: Bulgària, Croàcia, Xipre, Itàlia, Polònia, Romania, Suècia i Regne Unit. Per 12 d'ells el context és de desinterès polític: Àustria, Bèlgica, Estònia, Finlàndia, Grècia, Hongria Liechtenstein, Luxemburg, Lituània, Eslovàquia, Eslovènia i Espanya (aquesta indiferència política provoca que els organismes d'igualtat estiguin infrafinançats i sense tracció política per exercir les seves competències). Finalment, hi hauria un context de suport polític evident França, Alemanya, Irlanda, Islàndia, Letònia, Països Baixos i Portugal.

infrafinançament¹⁷¹⁰ o competències del personal inadequades¹⁷¹¹.

El potencial dels organismes d'igualtat dels estats membres per la promoció i la supervisió de l'aplicació de la normativa antidiscriminària, unit a les problemàtiques que s'han subratllat en relació amb la seva aplicació, van portar la Comissió Europea a publicar la Recomendación (UE) 2018/951 de la Comisión, de 22 de junio de 2018, sobre normas relativas a los organismos para la igualdad¹⁷¹² (d'ara en endavant, Recomanació 2018/951), amb la intenció d'incentivar tant l'ampliació de l'àmbit d'actuació dels organismes d'igualtat com les seves competències. Aquesta Recomanació va marcar l'inici del que en aquest treball anomenem una segona fase¹⁷¹³ en el desenvolupament dels organismes d'igualtat dels estats membres que, com veurem, acabarà tenint en compte la seva funció en relació amb la discriminació algorítmica.

1.2. Segona fase: ampliació de competències amb incidència directa en les facultats per fer front a la discriminació algorítmica

Aquesta segona fase en relació amb el desenvolupament dels organismes d'igualtat dels estats membres s'inicia amb la Recomanació 2018/951 (1.2.1), segueix amb l'Informe de la Comissió que avalua l'aplicació de la Recomanació de la Comissió 2018/951, de 22 de juny de 2018, de 19 de març de 2021 -d'ara en endavant, Informe- (1.2.2), on ja es fa referència a la supervisió de sistemes automatitzats, i culmina amb la presentació, el 7 de desembre de 2022, de dues propostes de Directiva (una que anomenarem proposta de Directiva sobre estàndards comuns pels organismes d'igualtat en l'àmbit de la igualtat de

¹⁷¹⁰ *Ídem, ibídem.*

¹⁷¹¹ Considerant 21 de la Recomendación (UE) 2018/951.

¹⁷¹² Recomendación (UE) 2018/951 de la Comisión, de 22 de junio de 2018, sobre normas relativas a los organismos para la igualdad. DOUE, L 167/28, 4.7.2018.

¹⁷¹³ REY MARTÍNEZ, F., "O organismo español para a igualdad...", *op. cit.*, 2021, pàg. 334. És en aquest article que l'autor parla d'una segona fase d'implementació dels organismes d'igualtat. L'autor estableix que amb la recomanació de 22 de juliol de 2018, sembla evident que la UE ha entrat en una segona fase d'implementació dels organismes per a la igualtat, després de la inicial obligació de creació a partir de l'any 2000 dels organismes, que no especificava, no obstant això, els trets fonamentals del seu règim d'organització i activitat. Aquesta segona fase es caracteritza pel disseny, a partir de la gradual fundació dels organismes en els diversos estats europeus i la seva experiència en els anys inicials de funcionament, d'un model d'organismes, tant precís com potent.

tractament entre homes i dones¹⁷¹⁴, i l'altra que anomenarem proposta de Directiva sobre estàndards comuns pels organismes d'igualtat en l'àmbit de la igualtat de tractament en relació amb l'origen ètnic o raça, creences o religió, edat, discapacitat o orientació sexual¹⁷¹⁵) per enfortir el rol dels organismes d'igualtat europeus, i que poden assegurar certes competències clau en el cas de la lluita envers la discriminació algorítmica (1.2.3).

1.2.1. Recomanació (UE) 2018/951, de 22 de juny de 2018, de la Comissió Europea

La Comissió Europea va adoptar la Recomanació 2018/951, per aconsellar l'establiment d'uns estàndards comuns pels organismes d'igualtat¹⁷¹⁶ amb la finalitat d'establir mesures per tal de millorar la seva independència i eficàcia, especialment en allò relatiu a la seva capacitat per garantir el ple gaudi dels drets per part de persones i grups víctimes de discriminació¹⁷¹⁷. En allò que fa referència a la protecció envers la discriminació algorítmica, és pertinent destacar diversos elements que es plantegen en la Recomanació 2018/951.

En primer lloc, es convida els estats membres a ampliar el mandat dels organismes d'igualtat per tal que s'inclogui la protecció envers la discriminació per motius de religió o conviccions, discapacitat, edat o orientació sexual¹⁷¹⁸, i que abasti tots els motius de

¹⁷¹⁴ EUROPEAN COMMISSION, *Proposal for a Council Directive on standards for equality bodies in the field of equal treatment between persons irrespective of their racial or ethnic origin, equal treatment in the field of employment and occupation between persons irrespective of their religion or belief, disability, age or sexual orientation, equal treatment between women and men in matters of social security and in the access to and supply of goods and services, and deleting Article 13 of Directive 2000/43/EC and Article 12 of Directive 2004/113/EC*, COM(2022) 689 final, Brussels, 7 December 2022.

¹⁷¹⁵ EUROPEAN COMMISSION, *Proposal for a Directive of the European Parliament and of the Council on standards for equality bodies in the field of equal treatment and equal opportunities between women and men in matters of employment and occupation, and deleting Article 20 of Directive 2006/54/EC and Article 11 of Directive 2010/41/EU*, COM(2022) 688 final, Brussels, 7 December 2022.

¹⁷¹⁶ EUROPEAN COMMISSION, *Equality bodies and the implementation...*, *op. cit.*, 19 May 2021, pàg. 3.

¹⁷¹⁷ Recomendación (UE) 2018/951, pàg. 32.

¹⁷¹⁸ En el Considerant 14 de la Recomendación (UE) 2018/951 es fa referència a la Propuesta de Directiva del Consejo por la que se aplica el principio de igualdad de trato entre las personas independientemente de su religión o convicciones, discapacidad, edad u orientación sexual, de 2 de juliol de 2008. Precisa que malgrat que la proposta no s'ha adoptat fins ara, convé encoratjar els estats membres a que designin organismes per a la igualtat en aquests àmbits (la protecció social, incloses la seguretat social i la sanitat, els avantatges socials, l'educació, l'accés a béns i serveis disponibles per al públic i l'oferta dels mateixos, inclòs l'habitatge) doncs l'experiència demostra que aquesta pràctica reforça la protecció envers la discriminació.

discriminació prohibits en els àmbits del treball (inclòs el treball autònom) i l'ocupació, l'accés a béns i serveis i el seu subministrament, l'educació, la protecció social i els beneficis socials, inclosa la incitació a l'odi relacionada amb aquests motius i dins d'aquests àmbits¹⁷¹⁹. La intenció patent de la Comissió Europea d'ampliar els àmbits d'actuació dels organismes d'igualtat fa afirmar a Rey Martínez que la garantia de l'organisme d'igualtat és un element permanent o sistèmic del quadre de garanties europeu en l'àmbit del dret antidiscriminatori i opera respecte als sis trets protegits pel dret comunitari¹⁷²⁰.

En segon lloc, s'estableixen diferents aspectes a tenir en compte en l'assistència independent a les víctimes, i contràriament a les directives, que no estableixen quines formes específiques d'assistència, la Recomanació 2018/951 sí que entra més en detall i inclou¹⁷²¹: rebre i tramitar denúncies individuals i col·lectives, proporcionar assessorament jurídic a les víctimes, també a l'hora de tramitar les seves denúncies, participar en activitats de mediació i conciliació, actuar en qualitat d'*amicus curiae* o d'expert on fos necessari, i com a aspecte més rellevant en relació amb la discriminació algorítmica, representar les persones denunciants davant dels tribunals. S'explicita en la Recomanació 2018/951 la necessitat de considerar l'opció d'oferir als organismes d'igualtat la possibilitat de participar o assistir a accions judicials, amb la finalitat de poder fer front a la discriminació estructural i sistèmica en els casos seleccionats pels propis organismes deguda la seva abundància, la seva gravetat o la necessitat de clarificació des del punt de vista jurídic, precisant que aquestes accions judicials poden tenir lloc tant en nom del propi organisme com en nom de les víctimes i organitzacions que les representen, de conformitat amb el dret processal nacional¹⁷²². La proposta d'objectivar la protecció contra les discriminacions, més enllà de la defensa subjectiva que s'ofereix a les víctimes, donant-los suport o impugnant amb legitimació processal activa les discriminacions estructurals en seu judicial, és a dir, la impugnació judicial directa per part de l'organisme¹⁷²³ és molt pertinent, en el cas de la discriminació

¹⁷¹⁹ Recomendación (UE) 2018/951, pàg. 32.

¹⁷²⁰ REY MARTÍNEZ, F., "O organismo español para a igualdad...", *op. cit.*, 2021, pàg. 317.

¹⁷²¹ EUROPEAN COMMISSION, *Equality bodies and the implementation...*, *op. cit.*, 19 May 2021, pàg. 7.

¹⁷²² Recomendación (UE) 2018/951, pàg. 32.

¹⁷²³ REY MARTÍNEZ, F., "O organismo español para a igualdad...", *op. cit.*, 2021, pàg. 329.

algorítmica¹⁷²⁴.

En tercer lloc, en el sentit d'incentivar l'actuació dels organismes d'igualtat en tant que òrgans quasi-jurisdiccionals, s'estableix que els estats membres han de possibilitar que els organismes d'igualtat recopilin proves i informació pertinents, de conformitat amb la legislació nacional, i han de considerar la possibilitat per part de l'organisme de formular recomanacions a l'hora de proporcionar assistència a les víctimes o quan la legislació nacional així ho autoritzi, decisions jurídicament vinculants en casos individuals o col·lectius de discriminació, així com un seguiment per garantir-ne l'aplicació. Si els organismes d'igualtat disposen de la capacitat legal necessària per adoptar decisions vinculants, els estats membres han de donar-los la capacitat també d'imposar sancions eficaces i proporcionades¹⁷²⁵. La capacitat de sancionar administrativament amb multes les infraccions tipificades que lesionin amb certa entitat la prohibició constitucional de discriminació, tenint oberta posteriorment, per part de la persona sancionada, la via de la protecció judicial corresponent, es pot equiparar als poders de les autoritats en matèria de protecció de dades, que arran de l'article 58.2 de l'RGPD han de disposar de poders coercitius, com per exemple, de sancionar i multar¹⁷²⁶.

Finalment, en quart lloc, també interessa que la Recomanació 2018/951 fa èmfasi en què els estats membres han de garantir que cada organisme disposi de recursos humans, tècnics i financers, així com dels locals i les infraestructures necessàries per desenvolupar les seves funcions i exercir les seves competències amb eficàcia, garantint que el

¹⁷²⁴ El fet que d'inici no es donés a l'organisme d'igualtat la funció de litigar al marge de les víctimes pot venir ocasionat per la concepció individualista de la discriminació, que dona molta més rellevància a la víctima de l'acció concreta discriminatòria (i la responsabilitza de tant l'inici com la recopilació de proves per l'acció judicial) i no al col·lectiu oprimint que finalment s'ha vist també perjudicat.

¹⁷²⁵ Recomanació (UE) 2018/951, pàg. 33; EUROPEAN COMMISSION, *Equality bodies and the implementation...*, *op. cit.*, 19 May 2021, pàg. 9. En l'Informe de la Comissió que avalua l'aplicació de la Recomanació 2018/951, de 22 de juny de 2018, de 19 de març de 2021, s'estableix que en alguns estats membres, els organismes d'igualtat tenen poder de decisió, incloent la capacitat de dur a terme recomanacions i decisions vinculants en casos de discriminació individual o col·lectiva. Si l'organisme d'igualtat té el poder de prendre decisions legalment vinculants en casos de discriminació, té la capacitat de posar sancions proporcionades efectives i adequades; estableix un recurs judicial efectiu i estàndards de procés degut, en particular quan es tracta de dret a apel·lar. En dos terços dels estats membres els organismes d'igualtat tenen el poder de dur a terme investigacions d'ofici. Malgrat que en la majoria d'estats membres els organismes d'igualtat poden fer recomanacions en les queixes de discriminació, només tenen el poder de prendre decisions legalment vinculants (és a dir, com un òrgan quasi-jurisdiccional) menys de 10. En la majoria d'aquests estats membres, els organismes d'igualtat poden defensar aquestes decisions davant un tribunal (si s'apel·la la seva decisió).

¹⁷²⁶ Article 58.2 de l'RGPD.

procediment per presentar denúncies és senzill i gratuït¹⁷²⁷.

Tenint en compte que la Recomanació 2018/951 és un acte jurídic de caràcter no obligatori i que era rellevant si els estats membres havien seguit o no les directrius que es proposaven, l'any 2021 es va publicar un informe de la Comissió avaluant-ne l'aplicació per part dels estats membres, que estudiem tot seguit.

1.2.2. Informe de la Comissió que avalua l'aplicació de la Recomanació (UE) 2018/951, de 22 de juny de 2018, de 19 de març de 2021

L'objectiu de l'Informe és fer balanç de la situació dels organismes d'igualtat en els estats membres de la UE establerts en l'article 13 de la Directiva 2000/43/CE, 12 de la Directiva 2004/113/CE, 20 de la Directiva 2006/54/CE i 11 de la Directiva 2010/41/UE i avaluar la implementació de la Recomanació 2018/951¹⁷²⁸.

Quant a la lluita envers la discriminació algorítmica, interessen diverses reflexions establertes en l'Informe. S'especifica que la possibilitat per part dels organismes d'igualtat per iniciar demandes en el seu nom¹⁷²⁹ està establerta en alguns estats membres¹⁷³⁰, incloent casos d'acció popular com també en algunes tradicions legals portar el cas d'una víctima identificable, amb el seu consentiment, però sense que la víctima formi part efectivament dels procediments judicials¹⁷³¹. S'explicita en l'Informe que això

¹⁷²⁷ Recomendación (UE) 2018/951, pàg. 34.

¹⁷²⁸ EUROPEAN COMMISSION, *Equality bodies and the implementation...*, op. cit., 19 May 2021, pàg. 2.

¹⁷²⁹ L'Informe es refereix a casos en els quals l'organisme d'igualtat porta un cas davant del tribunal sense una víctima identificable que sigui part del procediment.

¹⁷³⁰ EUROPEAN COMMISSION, *Equality bodies and the implementation...*, op. cit., 19 May 2021, pàg. 9. Tanmateix, també precisa que alguns estats membres han brindat poders molt limitats als organismes d'igualtat en relació amb la litigació, atès que no confereixen legitimació ni per representar a les víctimes davant dels tribunals, ni per iniciar procediments en el seu nom, ni per intervenir en suport d'una part o proporcionar informes com a *amicus curiae*. Entre els organismes d'igualtat que poden iniciar procediments legals, alguns van informar que havien portat més de 10 casos l'any 2019 en relació amb la discriminació per origen ètnic o racial, mentre que molts altres n'havien fet 0 o 1. Sembla ser que una raó per la inacció es degui a la manca de finançament combinada amb el risc d'haver de pagar costes quan es perd el cas. Vegeu: KÁDÁR, T., "The legal standing of equality bodies", *European equality law review*, núm. 1, 2019, pàg. 15. L'autor proposa que es revisi la norma segons la qual qui perd ha de pagar les costes en vista del desequilibri de poder que existeix en aquests casos.

¹⁷³¹ EUROPEAN COMMISSION, *Equality bodies and the implementation...*, op. cit., 19 May 2021, pàg. 7.

pot ser particularment beneficiós per abordar casos on la discriminació és causada per la intel·ligència artificial¹⁷³².

En el mateix sentit, l'Informe estableix que en casos de discriminació indirecta (específicament inclou aquí la discriminació causada per intel·ligència artificial), la recollida d'un conjunt complet, regular i comparable de dades sobre igualtat és de gran importància per permetre que els òrgans d'igualtat ofereixin assistència a les víctimes i contribueixin a l'aplicació efectiva dels drets¹⁷³³.

També sembla d'especial rellevància per la discriminació algorítmica que certs estats membres hagin desenvolupat relacions de treball amb les oficines nacionals estadístiques¹⁷³⁴ i la seva pròpia capacitat per fer-ne, malgrat s'alerta en l'Informe que l'infrafinançament limita la capacitat dels organismes d'igualtat per dur-les a terme¹⁷³⁵. A més, alguns organismes d'igualtat tenen com a funció fer proves, és a dir, dur a terme un mètode per evidenciar discriminació en base a la prova de comparació – per exemple, fer proves per un lloc de feina, amb candidatures idèntiques i només canviant característiques protegides de la població (p. e. informació relativa a l'edat)¹⁷³⁶. Aquesta funció és extrapolable al cas de la discriminació algorítmica, és a dir, fer proves amb el sistema algorítmic per detectar la possibilitat de resultats esbiaixats.

Quant a la funció dels organismes d'igualtat com a òrgans quasi-jurisdiccionals, l'Informe valora com a bona pràctica que l'organisme d'igualtat pugui imposar mesures cautelars, sancions administratives, ordenar el pagament d'una compensació incloent danys no econòmics, requerir a les persones infractores que cessin de la seva activitat i ordenar sancions si les seves exigències no es compleixen¹⁷³⁷. Així mateix, interessa també que en l'Informe es proposa com a bona pràctica dels organismes d'igualtat treballar en tres nivells, un general, buscant el progrés en totes les característiques protegides, un específic per a cada característica, i un tercer per casos de discriminacions múltiples o interseccionals, on interactuïn dos o més característiques¹⁷³⁸ (funció que analitzarem més

¹⁷³² *Ídem*, pàg. 8.

¹⁷³³ *Ídem*, pàg. 11. Tanmateix, en un terç dels estats membres els organismes d'igualtat no duen a terme enquestes ni ho han fet en molt de temps, establint com a causa principal la falta de recursos.

¹⁷³⁴ *Ídem*, pàg. 12.

¹⁷³⁵ *Ídem*, pàg. 20.

¹⁷³⁶ *Ídem*, pàg. 7 i 8.

¹⁷³⁷ *Ídem*, pàg. 10.

¹⁷³⁸ REY MARTÍNEZ, F., “O organismo español para a igualdad...” , *op. cit.*, 2021, pàg. 332.

específicament en l'apartat 2.3.1 d'aquest capítol).

L'Informe arriba a la conclusió que segueix essent un repte la infrarepresentació dels casos de discriminació; la manca de consciència per part de les víctimes dels seus drets i dels mecanismes existents per rebre assistència o reparació, inclús per part dels organismes d'igualtat¹⁷³⁹; o el fet que malgrat gaudeixin d'independència funcional, no sempre es garanteixi que siguin independents del govern a la pràctica¹⁷⁴⁰, ocasionant problemàtiques en contextos polítics poc consensuats¹⁷⁴¹. També es conclou que un nivell limitat i desigual d'implementació de la Recomanació 2018/951 continua obstaculitzant a alguns organismes d'igualtat dur a terme el seu rol i ocasiona diferents nivells de protecció contra la discriminació en la UE¹⁷⁴². Aquestes qüestions són la raó per la qual la Comissió Europea va presentar, el 7 de desembre de 2022, dues propostes de Directiva amb l'objectiu d'assegurar certes funcions als organismes d'igualtat dels estats membres de la UE.

1.2.3. Propostes de directives, de 7 de desembre de 2022, per establir normes comunes per als organismes d'igualtat nacionals

En base a la conclusió a la qual va arribar la Comissió Europea en l'Informe, va presentar el 7 de desembre de 2022 dues propostes de Directiva per enfortir el rol dels organismes d'igualtat europeus¹⁷⁴³. Considerem que aquest enfortiment pot tenir una repercussió directa en la lluita envers la discriminació algorítmica.

D'aquestes propostes cal destacar l'exigència als estats membres d'assegurar que els organismes d'igualtat disposin dels recursos humans, tècnics i financers necessaris per dur a terme totes les seves competències de manera efectiva, incloent l'eventual cas

¹⁷³⁹ EUROPEAN COMMISSION, *Equality bodies and the implementation...*, *op. cit.*, 19 May 2021, pàg. 15.

¹⁷⁴⁰ *Ídem*, pàg. 16.

¹⁷⁴¹ *Ídem*, pàg. 17-19. En l'informe es considera una bona pràctica per la independència dels organismes d'igualtat, que s'aplica en alguns estats membres, involucrar al Parlament -amb un vot de majoria qualificada. Inclús el Parlament pot ser responsable del nomenament d'una persona com a cap de l'organisme d'igualtat. També es considera bona pràctica que els organismes d'igualtat retin comptes a autoritats auditores.

¹⁷⁴² *Ídem*, pàg. 29.

¹⁷⁴³ CHOPIN, I. i GERMAINE, C., *A comparative analysis of...*, *op. cit.*, 2022, pàg. 107.

d'increment de competències, queixes, costos de litigació i l'ús de sistemes automatitzats¹⁷⁴⁴. També l'exigència d'assegurar que els organismes d'igualtat podran dur a terme investigacions¹⁷⁴⁵, i tindran drets efectius d'accés a la informació per establir que la discriminació ha ocorregut¹⁷⁴⁶, incloent la possibilitat que qui ha ocasionat la discriminació o una tercera part estiguin legalment obligats a donar la informació i els documents demanats pels organismes d'igualtat¹⁷⁴⁷ -en el cas que se'ls atorgués aquesta possibilitat, s'ha de puntualitzar que aquesta informació no es podria utilitzar posteriorment per iniciar procediments legals, sinó que només per expedir decisions pròpies de l'organisme d'igualtat¹⁷⁴⁸. A més, l'exigència d'assegurar que els organismes d'igualtat poden actuar en procediments legals¹⁷⁴⁹ com a part del procediment per defensar una decisió que han adoptat pròpiament¹⁷⁵⁰, aportant observacions com a *amicus curiae*¹⁷⁵¹ o iniciant o participant en els procediments en benefici o en suport de diverses

¹⁷⁴⁴ EUROPEAN COMMISSION, *Proposal for a Council Directive on standards for equality bodies in the field of equal treatment between persons irrespective of their racial or ethnic origin...*, *op. cit.*, 7 December 2022, art. 4; EUROPEAN COMMISSION, *Proposal for a Directive of the European Parliament and of the Council on standards for equality bodies in the field of equal treatment and equal opportunities between women and men...* " *op. cit.*, 7 December 2022, art. 4.

¹⁷⁴⁵ EUROPEAN COMMISSION, *Proposal for a Council Directive on standards for equality bodies in the field of equal treatment between persons irrespective of their racial or ethnic origin...*, *op. cit.*, 7 December 2022, art. 8.1; EUROPEAN COMMISSION, *Proposal for a Directive of the European Parliament and of the Council on standards for equality bodies in the field of equal treatment and equal opportunities between women and men...* " *op. cit.*, 7 December 2022, art. 8.1.

¹⁷⁴⁶ EUROPEAN COMMISSION, *Proposal for a Council Directive on standards for equality bodies in the field of equal treatment between persons irrespective of their racial or ethnic origin...*, *op. cit.*, 7 December 2022, art. 8.2; EUROPEAN COMMISSION, *Proposal for a Directive of the European Parliament and of the Council on standards for equality bodies in the field of equal treatment and equal opportunities between women and men...* " *op. cit.*, 7 December 2022, art. 8.2.

¹⁷⁴⁷ EUROPEAN COMMISSION, *Proposal for a Council Directive on standards for equality bodies in the field of equal treatment between persons irrespective of their racial or ethnic origin...*, *op. cit.*, 7 December 2022, art. 8.3; EUROPEAN COMMISSION, *Proposal for a Directive of the European Parliament and of the Council on standards for equality bodies in the field of equal treatment and equal opportunities between women and men...* " *op. cit.*, 7 December 2022, art. 8.3.

¹⁷⁴⁸ EUROPEAN COMMISSION, *Proposal for a Council Directive on standards for equality bodies in the field of equal treatment between persons irrespective of their racial or ethnic origin...*, *op. cit.*, 7 December 2022, art. 9.4 i considerant 33; EUROPEAN COMMISSION, *Proposal for a Directive of the European Parliament and of the Council on standards for equality bodies in the field of equal treatment and equal opportunities between women and men...* " *op. cit.*, 7 December 2022, art. 9.4 i considerant 32.

¹⁷⁴⁹ EUROPEAN COMMISSION, *Proposal for a Council Directive on standards for equality bodies in the field of equal treatment between persons irrespective of their racial or ethnic origin...*, *op. cit.*, 7 December 2022, art. 9.1; EUROPEAN COMMISSION, *Proposal for a Directive of the European Parliament and of the Council on standards for equality bodies in the field of equal treatment and equal opportunities between women and men...* " *op. cit.*, 7 December 2022, art. 9.1.

¹⁷⁵⁰ EUROPEAN COMMISSION, *Proposal for a Council Directive on standards for equality bodies in the field of equal treatment between persons irrespective of their racial or ethnic origin...*, *op. cit.*, 7 December 2022, art. 9.2.a; EUROPEAN COMMISSION, *Proposal for a Directive of the European Parliament and of the Council on standards for equality bodies in the field of equal treatment and equal opportunities between women and men...* " *op. cit.*, 7 December 2022, art. 9.2.a.

¹⁷⁵¹ EUROPEAN COMMISSION, *Proposal for a Council Directive on standards for equality bodies in the field of equal treatment between persons irrespective of their racial or ethnic origin...*, *op. cit.*, 7 December 2022, art. 9.2.b; EUROPEAN COMMISSION, *Proposal for a Directive of the European Parliament and*

víctimes amb la seva aprovació¹⁷⁵². També cal destacar l'exigència als estats membres d'assegurar que l'organisme d'igualtat pot actuar en el seu nom, en particular per adreçar discriminació estructural i sistemàtica en casos seleccionats per l'organisme d'igualtat per la seva abundància, seriositat o necessitat de clarificació legal¹⁷⁵³ (d'aquesta manera, la litigació estratègica serviria per promoure la lluita envers les discriminacions i provocar canvis legals i en les polítiques públiques¹⁷⁵⁴) sense alterar la legitimació de les víctimes, associacions, organitzacions i altres entitats legals¹⁷⁵⁵. De la mateixa manera, assegurar que els organismes d'igualtat tenen mecanismes per cooperar amb altres organismes tant nacionals (com podrien ser les autoritats de protecció de dades o de la intel·ligència artificial), com de la UE o internacionals¹⁷⁵⁶. I endemés, s'especifica que els organismes d'igualtat han de tenir accés a estadístiques relacionades amb els drets d'igualtat i no-discriminació fetes tant per entitats privades com públiques per avaluar la situació de discriminació en l'estat membre¹⁷⁵⁷.

of the Council on standards for equality bodies in the field of equal treatment and equal opportunities between women and men... " *op. cit.*, 7 December 2022, art. 9.2.b.

¹⁷⁵² EUROPEAN COMMISSION, *Proposal for a Council Directive on standards for equality bodies in the field of equal treatment between persons irrespective of their racial or ethnic origin...*, *op. cit.*, 7 December 2022, art. 9.2.c; EUROPEAN COMMISSION, *Proposal for a Directive of the European Parliament and of the Council on standards for equality bodies in the field of equal treatment and equal opportunities between women and men...* " *op. cit.*, 7 December 2022, art. 9.2.c.

¹⁷⁵³ EUROPEAN COMMISSION, *Proposal for a Council Directive on standards for equality bodies in the field of equal treatment between persons irrespective of their racial or ethnic origin...*, *op. cit.*, 7 December 2022, art. 9.3; EUROPEAN COMMISSION, *Proposal for a Directive of the European Parliament and of the Council on standards for equality bodies in the field of equal treatment and equal opportunities between women and men...* " *op. cit.*, 7 December 2022, art. 9.3.

¹⁷⁵⁴ REY MARTÍNEZ, F., "O organismo español para a igualdad...", *op. cit.*, 2021, pàg. 332. Per una informació detallada sobre la litigació estratègica, vegeu: EUROPEAN NETWORK OF EQUALITY BODIES (EQUINET), *Strategic Litigation Handbook*, Brussels, 2017. Per una defensa de l'ús de la litigació estratègica per part de sindicats per protegir i reforçar els drets laborals individuals que puguin veure's compromesos per la gestió algorítmica, vegeu: DE STEFANO, V. i TAES, S., "Algorithmic management and collective bargaining", *Transfer: European Review of Labour and Research*, vol. 29, núm. 1, 2023, pàg. 30.

¹⁷⁵⁵ EUROPEAN COMMISSION, *Proposal for a Council Directive on standards for equality bodies in the field of equal treatment between persons irrespective of their racial or ethnic origin...*, *op. cit.*, 7 December 2022, considerant 35; EUROPEAN COMMISSION, *Proposal for a Directive of the European Parliament and of the Council on standards for equality bodies in the field of equal treatment and equal opportunities between women and men...* " *op. cit.*, 7 December 2022, considerant 34.

¹⁷⁵⁶ EUROPEAN COMMISSION, *Proposal for a Council Directive on standards for equality bodies in the field of equal treatment between persons irrespective of their racial or ethnic origin...*, *op. cit.*, 7 December 2022, art. 12; EUROPEAN COMMISSION, *Proposal for a Directive of the European Parliament and of the Council on standards for equality bodies in the field of equal treatment and equal opportunities between women and men...* " *op. cit.*, 7 December 2022, art. 12.

¹⁷⁵⁷ EUROPEAN COMMISSION, *Proposal for a Council Directive on standards for equality bodies in the field of equal treatment between persons irrespective of their racial or ethnic origin...*, *op. cit.*, 7 December 2022, art. 14.2; EUROPEAN COMMISSION, *Proposal for a Directive of the European Parliament and of the Council on standards for equality bodies in the field of equal treatment and equal opportunities between women and men...* " *op. cit.*, 7 December 2022, art. 14.2.

Cal puntualitzar que les regulacions de la UE en l'àmbit dels organismes d'igualtat són, o bé no obligatòries (com la Recomanació 2018/951) o bé regulacions de mínims (Directiva 2000/43/CE, Directiva 2004/113/CE, Directiva 2006/54/CE, Directiva 2010/41/UE), de manera que els estats membres poden dur a terme regulacions més protectores, donant un àmbit d'actuació més extens (tant afegint característiques protegides com àmbits materials d'aplicació), donant més competències o més poders.

A tall de recapitulació, la normativa de la UE que regula els organismes d'igualtat està evolucionant en el sentit de donar-los un rol cada cop més rellevant i inclús fonamental en l'aplicació de la normativa antidiscriminació. Creiem que tant en el cas de la discriminació algorítmica -on la problemàtica deguda la maca d'informació es fa més evident-, com també en la discriminació tradicional¹⁷⁵⁸, la intenció de la UE és assegurar un rol proactiu per part d'organismes que vetllen per l'interès general en matèria d'igualtat de tractament, atès que a les persones que es veuen discriminades els és molt difícil defensar-se (en el cas específic de la discriminació algorítmica, desafiar la decisió presa pel sistema). Tot seguit, estudiem les funcions que sota la nostra perspectiva ha de dur a terme l'organisme d'igualtat dels estats membres de la UE per proporcionar una millor protecció i configurar-se com una eina útil i efectiva en la lluita envers la discriminació algorítmica.

2. Funcions de l'organisme d'igualtat per fer front a la discriminació algorítmica

En aquest apartat concretem les funcions que poden dur a terme els organismes nacionals dels estats membres per ajudar proactivament a les persones i grups perjudicats en casos en què s'ha produït discriminació algorítmica. L'atribució d'aquestes funcions als organismes d'igualtat ve determinada per l'experiència, que mostra, inclús en la discriminació tradicional, que molt poques persones que s'han sentit víctimes d'una discriminació plantegen demandes per si soles, presumiblement perquè l'acció judicial és incomprendible per a elles, estressant, dilatada en el temps i sense clares expectatives

¹⁷⁵⁸ CHOPIN, I. i GERMAINE, C., A comparative analysis of..., *op. cit.*, 2022, pàg. 131. Les autores destaquen el rol dels organismes d'igualtat per reduir la discrepància entre els nivells de discriminació experimentada i aquella que es denuncia.

d'èxit¹⁷⁵⁹. Tanmateix, per dur a terme aquestes funcions és essencial que l'organisme d'igualtat sigui independent de qualsevol part, és a dir, que se sotmeti a requeriments de transparència per evidenciar que la seva tasca és en favor de l'interès general¹⁷⁶⁰, i també disposi del nivell de finançament adequat per tenir un impacte real¹⁷⁶¹.

Tenint en compte la normativa i les recomanacions establertes per part de la UE, hem organitzat l'apartat entre les funcions per part de l'organisme d'igualtat per fer front a la discriminació algorítmica com a òrgan quasi-jurisdiccional (2.1), com a òrgan amb legitimació per iniciar procediments en matèria de discriminació (2.2), i altres funcions que poden complementar la seva tasca com a referent en la lluita envers la discriminació algorítmica (2.3).

2.1. Funcions com a òrgan quasi-jurisdiccional

En aquest apartat estudiem, en primer lloc, la possibilitat per part dels organismes d'igualtat de dur a terme investigacions pròpies (2.1.1), i en segon lloc, de fer, segons el resultat de les investigacions, decisions -vinculants o no-, i imposar sancions (2.1.2).

2.1.1. Possibilitat d'iniciar investigacions pròpies

La possibilitat d'iniciar investigacions pròpies per part de l'organisme d'igualtat dels estats membres pot millorar ostensiblement la protecció envers la discriminació algorítmica. Ja sigui arran de la informació obtinguda per part de l'organisme pròpiament, arran de la queixa d'una persona, o arran de la informació que pot obtenir a través de la documentació tècnica de sistemes d'IA de risc alt, podrà considerar l'existència d'indicis de discriminació, indagant i profunditzant sobre el cas en qüestió.

Des de la normativa europea es considera una millor pràctica que l'organisme d'igualtat

¹⁷⁵⁹ REY MARTÍNEZ, F., “O organismo español para a igualdad...”, *op. cit.*, 2021, pàg. 326 i 327.

¹⁷⁶⁰ HODGES, C., “Collective Redress: The Need for New Technologies”, *op. cit.*, 2019, pàg. 81.

¹⁷⁶¹ CROWLEY, N., *Equality bodies making a difference*, *op. cit.*, 2018, pàg. 11.

tingui la potestat d'escollar i investigar reclamacions de discriminació. És a dir, en comptes de limitar-se a proveir informació a la víctima sobre les normes i els seus drets, derivar-la directament a una organització civil o ajudar-la a assolir un acord amistós, tenir la capacitat d'investigar i resoldre queixes i reclamacions, i també dictar resolucions (vinculants o no)¹⁷⁶². L'abast de la potestat investigadora varia d'uns països als altres, des de vagues provisions (com a Itàlia) a situacions en les quals l'organisme té una posició semblant a la d'un fiscal o un jutge d'instrucció (com a França)¹⁷⁶³. En alguns casos té la potestat de sancionar a qui es negui a proporcionar la informació sol·licitada (com a Hongria¹⁷⁶⁴ o a Espanya), malgrat que, en aquest últim cas, l'organisme d'igualtat no podrà pròpiament imposar la sanció sinó que haurà d'interessar l'actuació de l'Administració de l'Estat (vegeu l'apartat 3.2 d'aquest capítol).

Aquesta possibilitat va lligada específicament a que els organismes d'igualtat disposin dels recursos humans, tècnics i financers necessaris per dur a terme totes les seves competències de manera efectiva, incloent si aquestes competències impliquen fer front a la discriminació causada per l'ús de sistemes automatitzats. Per tant, la facultat de dur a terme investigacions pròpies en l'àmbit de la discriminació algorítmica està estretament lligada a la disponibilitat de recursos per abordar com s'ocasiona aquesta discriminació i quins són els elements que han de ser objecte d'investigació.

Finalment, cal subratllar que un organisme d'igualtat amb capacitat per fer investigacions pròpies i capacitat de fer resolucions vinculants (amb possibilitat de revisió judicial), agilitzaria la resolució de casos de discriminació algorítmica en comparació amb les investigacions que es duen a terme en el si d'un procediment judicial per exemple, si l'organisme d'igualtat iniciés una acció de cessació d'un sistema algorítmic discriminatori. Tanmateix, les resolucions vinculants de l'organisme d'igualtat haurien de poder exigir la cessació del sistema algorítmic discriminatori, perquè de no ser així la seva actuació no seria tant efectiva com l'acció judicial. Per tant, seguidament analitzem la importància que l'organisme d'igualtat pugui fer resolucions i imposar sancions en el cas de discriminació algorítmica.

¹⁷⁶² REY MARTÍNEZ, F., "O organismo español para a igualdad...", *op. cit.*, 2021, pàg. 326 i 327.

¹⁷⁶³ *Ídem*, pàg. 327.

¹⁷⁶⁴ *Ídem, ibídem*.

2.1.2. Possibilitat de prendre decisions vinculants i imposar sancions si hi ha violacions del principi de no-discriminació

Com ja hem esmentat, és tan crucial pels organismes d'igualtat tenir la possibilitat de realitzar investigacions com prendre decisions vinculants i imposar sancions per incompliment¹⁷⁶⁵. Hodges fa referència, com a exemple d'òrgan administratiu sancionador, al Wettbewerbszentrale (WBZ) alemany, un òrgan independent i intersectorial encarregat de supervisar l'equitat en la competició des de 1912, que basa el seu èxit en el seu compromís amb l'interès públic, en protegir el procés de competència i el funcionament òptim del mercat i l'interès general, no la suma d'interessos individuals dels participants en el mercat¹⁷⁶⁶. L'autor planteja que un gran avantatge d'aquest sistema radica en què no s'ha de divulgar qui ha alertat l'organisme en qüestió de la infracció o a qui afavorirà la cessació de la infracció, i per tant, l'anonimat pot assegurar-se i no hi ha risc que la pressió per part de grans corporacions previngui l'interès general de ser protegit¹⁷⁶⁷. També, per exemple, els organismes de defensoria dels drets d'igualtat (Equality ombudsperson) que poden investigar queixes en relació amb la discriminació i dependent de la seva competència, poden prendre decisions o usar mesures més toves per alterar les pràctiques discriminatòries¹⁷⁶⁸. Tanmateix, cal fer esment al treball de Chopin i Germaine, segons el qual dels 44 organismes d'igualtat europeus¹⁷⁶⁹ estudiats, només 19 tenen funcions quasi-jurisdictionals i d'aquests, només 13 poden emetre decisions vinculants¹⁷⁷⁰.

En comparació, les autoritats independents dels estats membres plantejades en relació amb la protecció de dades (en el cas espanyol, l'AEPD), sí que actuen com a òrgans quasi-jurisdictionals operatius i, des de la normativa de la UE, s'exigeix que les autoritats puguin dur a terme aquestes tasques. L'RGPD determina que en cada estat membre de la UE hauran d'existir una o varies autoritats de control, independents dels restants poders públics, constituïdes conforme al dret públic per dur a terme les funcions que l'RGPD li atribueix, que puguin interaccionar amb les autoritats comunitàries i les autoritats de

¹⁷⁶⁵ ALLEN, R. i MASTERS, D., *Regulating for an equal AI...*, *op. cit.*, 2020, pàg. 68.

¹⁷⁶⁶ HODGES, C., "Collective Redress: The Need for New Technologies", *op. cit.*, 2019, pàg. 82. Per més informació sobre aquest òrgan, vegeu: <https://www.wettbewerbszentrale.de/de/home/> Darrera visita el 2 de setembre de 2023.

¹⁷⁶⁷ *Idem*, pàg. 83.

¹⁷⁶⁸ HAKKARAINEN, J., "Naming something collective...", *op. cit.*, 2021, pàg. 16.

¹⁷⁶⁹ Ens referim aquí a "europeus" en comptes de "de la UE", perquè malgrat s'integren tots els estats membres de la UE en l'estudi també s'inclouen altres estats europeus que no són de la UE.

¹⁷⁷⁰ CHOPIN, I. i GERMAINE, C., A comparative analysis of..., *op. cit.*, 2022, pàg. 120 i 121.

control dels altres estats membres¹⁷⁷¹. L'Estatut de l'autoritat de control de cada estat membre es troba comprès en l'RGPD, corresponent a la legislació de cada país regular qüestions menors o de detall, que en cap cas podran menystenir la capacitat de l'autoritat de control per dur a terme la seva tasca (per la qual ha de comptar amb suficients recursos¹⁷⁷²), precisant en el mateix RGPD els poders coercitius que hauran de tenir totes les autoritats de control de protecció de dades (entre ells, poders per sancionar)¹⁷⁷³.

Per tant, creiem que per fer efectives les decisions que pugui prendre l'organisme d'igualtat en tant que òrgan quasi-jurisdiccional, és essencial que l'òrgan tingui potestat sancionadora. No només per l'efectivitat sinó també per la visibilitat de l'òrgan en tant que protector dels drets de les persones enfront la discriminació algorítmica. Tal com planteja Rey Martínez, la capacitat de sancionar administrativament amb multes les infraccions tipificades que lesionin amb certa entitat la prohibició constitucional de discriminació, tenint oberta posteriorment per la part sancionada la via de la protecció judicial corresponent, donaria a la institució més visibilitat davant la població¹⁷⁷⁴.

Posem un exemple d'una actuació òptima d'un organisme d'igualtat en tant que òrgan quasi-jurisdiccional. Una persona que creu haver-se vist discriminada mitjançant un sistema algorítmic presenta una queixa davant l'organisme d'igualtat (també seria possible que d'ofici l'organisme d'igualtat iniciés la investigació) sobre l'actuació d'un sistema algorítmic determinat, per exemple, un sistema algorítmic de contractació d'una gran empresa o un sistema algorítmic de gestió d'una ajuda pública. Aquest, a través de les seves competències d'investigació podrà demanar a l'empresa o organització o administració pública, que presumptivament disposa d'un sistema algorítmic causant de discriminació, que li proporcioni informació sobre la precisió del sistema algorítmic segons determinats grups de població. L'empresa, organització o administració pública, si l'estat membre ho ha disposat així, estarà obligada a donar aquesta informació, ja que si no ho fa, l'organisme podrà sancionar-la. Si s'acredita que el sistema algorítmic no és igual de precís segons el grup de població al qual s'apliqui, l'organisme d'igualtat podrà

¹⁷⁷¹ MORENO MONTES DE OCA, B., "Autoridad de Control (Comentario al artículo 4.21 RGPD)", a TRONCOSO REIGADA, A., (dir.) i VILLALBA CANO, L., (coord.), *Comentario al Reglamento General de Protección de Datos y a la Ley Orgánica de Protección de Datos Personales y Garantía de los Derechos Digitales*, Tomo I, Pamplona, Aranzadi, 2021, pàg. 805.

¹⁷⁷² En relació amb l'escassetat de recursos de les autoritats de protecció de dades, vegeu: KAMINSKI, M.E., "Binary Governance...", *op. cit.*, 2019, pàg. 1599.

¹⁷⁷³ Article 58.2 de l'RGPD.

¹⁷⁷⁴ REY MARTÍNEZ, F., "O organismo español para a igualdad...", *op. cit.*, 2021, pàg. 329 i 339.

imposar una sanció i prendre la decisió que s'ha de cessar l'ús del sistema algorítmic que està causant aquesta discriminació. Aquesta decisió haurà de ser susceptible de recurs a instància judicial. És rellevant precisar que l'organisme d'igualtat ha de tenir la competència d'emetre decisions vinculants i no únicament posar sancions, ja que sancionar l'empresa, organització o administració pública per l'ús d'un sistema algorítmic discriminatori sense obligar-la a cessar l'ús del sistema pot no tenir efectivitat en el sentit de protegir les persones envers al discriminació algorítmica.

En el mateix sentit, en relació amb l'ús de mètodes alternatius de solució de conflictes¹⁷⁷⁵ en l'àmbit de la discriminació algorítmica, no considerem que puguin ser de molta utilitat, ja que arribar a un acord entre parts (entre les quals hi ha un desequilibri de poder evident que dificulta molt assolir un pacte just i adequat) que no impliqui o bé el cessament de l'ús del sistema algorítmic, o bé una disculpa pública, no és útil per afrontar el problema d'arrel, és a dir, un sistema d'opressió, sinó que l'objectiu acaba essent únicament la satisfacció individual de les parts i no té efectes més enllà d'aquestes.

Estudiem, a continuació, la necessitat que l'organisme d'igualtat pugui iniciar accions judicials per protegir les persones envers la discriminació algorítmica.

2.2. Funcions com a òrgan legitimat per iniciar litigis per protegir persones i col·lectius discriminats

Per dur a terme la seva funció de manera efectiva com a òrgan legitimat per iniciar litigis per protegir persones i col·lectius discriminats, l'organisme d'igualtat ha de poder tenir accés a informació (2.2.1), fer proactivament estadístiques que serveixin com a element de comparació (2.2.2), gestionar una plataforma que interrelacioni persones perjudicades

¹⁷⁷⁵ HODGES, C., "Collective Redress: The Need for New Technologies", *op. cit.*, 2019, pàg. 68. Segons l'autor, el desenvolupament per part de defensories del poble de tècniques de resolució de disputes com a alternativa a la litigació en l'àmbit del consum (que també pot incloure la discriminació), només s'empra en certes jurisdiccions, però habitualment opera en sectors de mercat específics com els serveis financers, energètics, de comunicació i de correus. Alguns ombudsperson operen en un model d'arbitratge (com en les assegurances, transport i energia a Alemanya) però altres han evolucionat per desenvolupar tècniques de resolució de disputes. El model ideal inclouria les següents tècniques: donar informació imparcial a persones consumidores, avaluar casos com a triatge, establir queixes formals, obtenir informació de cada part del conflicte, explorar oportunitats per les parts per accedir a un pacte, prendre alguna decisió en algun cas (en els esquemes es varia sobre si la decisió és una recomanació o legalment vinculant).

(2.2.3), per finalment tenir suficients indicis per poder litigar i sol·licitar al tribunal tant la cessació d'un sistema algorítmic que estigui causant discriminació com la indemnització per a les persones afectades (2.2.4).

2.2.1. Possibilitat d'accés a informació

Com hem vist al llarg d'aquest treball, no hi ha moltes disposicions normatives que possibilitin accedir a les persones afectades a informació relativa a una decisió presa mitjançant sistemes algorítmics (únicament l'article 13.2.f, 14.2.g, i 15.1.h de l'RGPD). Aquest accés limitat a la informació no sembla ser suficient per desafiar la discriminació algorítmica, per això ens decantem per considerar que a falta d'exigència de transparència enfront el públic general, són els organismes públics els que, si més no, han de tenir accés a certa informació sobre els sistemes algorítmics, no només per fer un control de transparència sinó per tenir informació per poder desafiar les decisions quan es produeix discriminació algorítmica. Tanmateix, un dels problemes per accedir a informació de rellevància quan ens referim a sistemes algorítmics complexos i opacs, que ja s'ha detectat en relació amb les autoritats de protecció de dades, és que molts estats membres no tenen els recursos per assumir aquesta tasca¹⁷⁷⁶.

Tal com estableix la Recomanació 2018/951, és imprescindible que els estats membres tinguin accés a informació de les empreses, organitzacions o administracions públiques que presumptivament han causat l'acció discriminatòria i inclús, en el cas que aquestes es neguin a donar la informació, imposar sancions. Com hem referenciat en el cas de realitzar investigacions pròpies en tant que òrgan quasi-jurisdiccional, seria pertinent que l'organisme d'igualtat pogués accedir a la informació. Tot i així, cal tenir en compte que aquesta informació obtinguda en tant que òrgan quasi-jurisdiccional, posteriorment - excepte que es vulgui defensar una decisió presa per l'òrgan que ha estat objecte de recurs judicial o en la seva actuació com a *amicus curiae*-, no es podrà utilitzar. Els estats membres s'han d'assegurar que en tant que òrgan quasi-jurisdiccional, l'organisme d'igualtat té eines per accedir a certa informació de rellevància per la detecció de la

¹⁷⁷⁶CASTETS-RENARD, C., "Accountability of Algorithms in the GDPR...", *op. cit.*, 2019, pàg. 125.

discriminació algorítmica -és a dir, resultats donats pel sistema, precisió del sistema segons grups determinats de població, proves que s'hagin fet al sistema per avaluar-ne la conformitat- i que, en tant que òrgan legitimat per iniciar litigis per discriminació algorítmica, l'organisme d'igualtat pugui demanar al tribunal que l'altra part reveli informació que té en la seva possessió i a la qual les persones afectades no tenen accés.

En aquest sentit cal referenciar l'article 18 de la Directiva 2020/1828, que obliga els estats membres a vetllar per tal que quan una entitat habilitada (que podria ser l'organisme d'igualtat en matèria de discriminació algorítmica) hagi aportat proves raonablement disponibles, suficients per donar suport a una acció de representació, i hagi indicat altres proves que estiguin en el poder de la part demandada o d'una tercera part, l'òrgan jurisdiccional o l'autoritat administrativa pugui ordenar a la part demandada o tercera part que exhibeixi les proves sol·licitades¹⁷⁷⁷ (vegeu apartat 3.3. del capítol VII). També estava redactada en aquest sentit l'article 4.b de la Propuesta de Directiva del Consejo relativa a la carga de la prueba en el ámbito de la igualdad de retribución y trato entre hombres y mujeres, on s'establia que les parts en litigi podien disposar de qualsevol informació necessària per fer valer els seus drets en poder de l'altra part¹⁷⁷⁸ (vegeu apartat 4.2 del capítol IV).

2.2.2. Possibilitat de confecció d'estadístiques

La Recomanació 2018/951 estableix que els estats membres haurien de permetre als organismes d'igualtat la realització d'estudis independents, i assegurar que puguin recopilar una quantitat suficient de dades fiables, quantitatives i qualitatives sobre la discriminació que possibiliti l'anàlisi necessari per extreure'n conclusions¹⁷⁷⁹. En l'Informe de la Comissió que avalua l'aplicació de la Recomanació 2018/951, es destaca com a bona pràctica el desenvolupament de relacions de treball amb les oficines nacionals estadístiques¹⁷⁸⁰ i la pròpia capacitat dels organismes d'igualtat per fer-ne, com també en

¹⁷⁷⁷ Article 18 de la Directiva 2020/1828.

¹⁷⁷⁸ COMISIÓN DE LAS COMUNIDADES EUROPEAS, *Propuesta de Directiva del Consejo relativa a la carga de la prueba en el ámbito de la igualdad...*, op. cit., 27 de mayo de 1988, art. 4.b.

¹⁷⁷⁹ Recomendación (UE) 2018/951, pàg. 33.

¹⁷⁸⁰ EUROPEAN COMMISSION, EUROPEAN COMMISSION, *Equality bodies and the implementation...*, op. cit., 19 May 2021, pàg. 12.

les propostes de directives sobre estàndards comuns dels organismes d'igualtat, de 7 de desembre de 2022, s'especifica que els organismes d'igualtat han de tenir accés a estadístiques relacionades amb els drets d'igualtat i no-discriminació fetes tant per entitats privades com públiques per avaluar la situació de discriminació en l'estat membre¹⁷⁸¹.

La posició dels organismes d'igualtat dels estats membres com a detentors i contribuïdors d'informació sobre les dades relatives a la discriminació en l'estat membre en qüestió genera la possibilitat que puguin emprar aquesta informació com a criteri comparatiu en el cas de la discriminació algorítmica. Com hem esmentat en l'apartat 3.3.2 del capítol II, estadístiques en relació amb assumptes complexos moltes vegades directament no estan disponibles¹⁷⁸², fet que dificulta la prova de discriminació algorítmica¹⁷⁸³.

En aquest sentit, les Guies per millorar la recollida i l'ús de dades d'igualtat del Grup d'Alt Nivell en antidiscriminació, igualtat i diversitat¹⁷⁸⁴, donen certs criteris rellevants per a la confecció d'estadístiques per part d'organismes, tal com els organismes d'igualtat, que poden tenir una incidència majúscula en la lluita envers la discriminació algorítmica. Les dades d'igualtat¹⁷⁸⁵ es defineixen com qualsevol tipus d'informació (tant qualitativa com quantitativa) que és útil per descriure i analitzar la igualtat en un context determinat¹⁷⁸⁶. Hi ha diferents fonts de dades i mecanismes de recollida que poden emprar-se per confeccionar estadístiques, com per exemple, el cens de població, registres administratius, enquestes, o dades sobre queixes de discriminació, entre altres¹⁷⁸⁷. En aquestes guies s'estableix que si les dades personals desagregades per característiques protegides com la raça o l'origen ètnic (que per raons històriques tendeixen a estar molt protegides), es recullen i es processen amb total respecte pel marc legal i les salvaguardes

¹⁷⁸¹ EUROPEAN COMMISSION, *Proposal for a Council Directive on standards for equality bodies in the field of equal treatment between persons irrespective of their racial or ethnic origin...*, *op. cit.*, 7 December 2022, art. 14.2; EUROPEAN COMMISSION, *Proposal for a Directive of the European Parliament and of the Council on standards for equality bodies in the field of equal treatment and equal opportunities between women and men...* *op. cit.*, 7 December 2022, art. 14.2.

¹⁷⁸² WACHTER, S., MITTELSTADT, B. i RUSSELL, C., "Why fairness cannot be automated...", *op. cit.*, 2021, pàg. 34.

¹⁷⁸³ ORGANISATION FOR SECURITY AND CO-OPERATION IN EUROPE (OSCE), *Use of actio popularis...*, *op. cit.*, November 2016, pàg. 13 i 14.

¹⁷⁸⁴ HIGH LEVEL GROUP ON NON-DISCRIMINATION, EQUALITY, AND DIVERSITY. SUBGROUP ON EQUALITY DATA, *Guidelines on improving the collection and use of equality data*, Brussels, 2018.

¹⁷⁸⁵ En anglès, "equality data".

¹⁷⁸⁶ HIGH LEVEL GROUP ON NON-DISCRIMINATION, EQUALITY, AND DIVERSITY. SUBGROUP ON EQUALITY DATA, *Guidelines on improving the collection and use of equality data*, Brussels, 2018, pàg. 8.

¹⁷⁸⁷ *Idem*, pàg. 8.

establertes, poden ser essencials¹⁷⁸⁸ tant per avaluar les situacions de discriminació en l'estat membre com per l'ús d'aquestes dades amb finalitats comparatives en casos de discriminació indirecta. La recollida i l'ús de dades d'igualtat poden ser indispensables, entre altres coses, per fer visible les formes amagades de discriminació i proveir evidències confiablès en casos judicials o administratius a través de dades que apunten cap a discriminació directa o indirecta¹⁷⁸⁹.

Actualment, segons les guies, hi ha certs buits i desafiaments comuns en tots els estats membres, com per exemple, i entre altres, la manca d'una perspectiva coordinada entre els organismes que recullen i usen les dades d'igualtat, la manca de recursos suficients i sensibilització sobre la importància de recollir dades d'igualtat, el desequilibri entre la recollida de dades sobre diferents característiques protegides (que pot ocasionar que manquin evidències sobre fins a quin punt les persones pateixen discriminació múltiple), el desequilibri entre la recollida de dades segons les diferents àrees, i la interpretació poc acurada de la normativa de protecció de dades (l'article 9.2.j de l'RGPD possibilita el tractament de categories especials de dades si és necessari per finalitats estadístiques¹⁷⁹⁰, i el considerant 26 de l'RGPD¹⁷⁹¹ estableix que els principis de protecció de dades no s'apliquen a la informació anònima, incloent per finalitats estadístiques)¹⁷⁹².

En les guies s'estableix que es podria donar un mandat específic, amb el finançament adequat¹⁷⁹³, als organismes d'igualtat (com també als organismes dels estats membres o a un institut de recerca) que identifiqués diferents fonts i proveïdors de dades d'igualtat, identifiqués sobre quines característiques protegides es recullen dades, que tingués en compte les interseccions entre les característiques protegides, identifiqués en quines àrees de la vida les dades es recullen (per exemple, accés a l'habitatge, a l'educació, a la sanitat, a la justícia, a béns i serveis, etc.), recollís informació sobre la representativitat (incloent com certs grups sistemàticament són infrarepresentats o exclosos) i la grandària de la base de dades, o establís amb quina freqüència i en base a quines definicions i categoritzacions

¹⁷⁸⁸ *Ídem*, pàg. 7.

¹⁷⁸⁹ *Ídem*, pàg. 9.

¹⁷⁹⁰ Article 9.2.j de l'RGPD.

¹⁷⁹¹ Considerant 26 de l'RGPD.

¹⁷⁹² HIGH LEVEL GROUP ON NON-DISCRIMINATION, EQUALITY, AND DIVERSITY. SUBGROUP ON EQUALITY DATA, *Guidelines on improving the collection and use of equality data*, Brussels, 2018, pàg. 9-12.

¹⁷⁹³ *Ídem*, pàg. 16.

es recullen les dades¹⁷⁹⁴.

Per exemple, per incorporar informació sobre la discriminació, entre altres, s'hauria de sistematitzar la introducció de variables i qüestions en les enquestes que mesuressin les desigualtats i la percepció en un nombre determinat de característiques, desagregar dades recollides -fins al punt on sigui possible- en base a l'edat, el gènere, la raça o l'origen ètnic, la migració, l'orientació sexual o la discapacitat, donar suport al desenvolupament i la implementació d'enquestes que puguin captar millor grups de població dispersos i de difícil accés en risc de discriminació o identificar àrees geogràfiques on les desigualtats s'observin més sovint¹⁷⁹⁵.

La capacitat dels organismes d'igualtat per obtenir i generar estadístiques en relació amb la discriminació en un estat membre pot facultar-los per emprar aquestes estadístiques amb la finalitat de centrar la seva atenció en els àmbits en què més incidència té la discriminació algorítmica i també per, en els casos en els quals participin com a part, presentar proves que serveixin com a element de comparació.

2.2.3. Gestió d'una plataforma que interrelacioni persones perjudicades

Arran de l'estudi de l'accés a la prova de discriminació algorítmica que s'ha fet al llarg del treball, i sobretot de l'impacte de la discriminació en més d'una persona a la vegada per part del mateix sistema algorítmic, creiem rellevant que existeixi una plataforma pública i gestionada per part d'un únic organisme -el millor posicionat és l'organisme d'igualtat dels estats membres-, que centralitzi les queixes que existeixin sobre discriminació patida arran de sistemes algorítmics i pugui, a través de la informació obtinguda, tant relacionar les persones que s'hagin pogut veure discriminades per un mateix sistema algorítmic, com investigar si s'ha produït la discriminació, iniciant accions de cessació perquè l'empresa, organització o administració pública que l'empri estigui obligada a deixar de fer-ho. Per tant, la interrelació de les persones que es trobin

¹⁷⁹⁴ *Ídem*, pàg. 13 i 14. Les guies estableixen la necessitat d'assegurar la comparabilitat de les dades d'igualtat que es veu fortament afectada quan els conceptes, definicions o categories difereixen segons les fonts de dades, el moment temporal o els països, i es pot minimitzar, per exemple, treballant en identificar les diferències en les definicions o harmonitzant-les (pàg. 20 i 21).

¹⁷⁹⁵ *Ídem*, pàg. 18 i 19.

discriminades mitjançant sistemes algorítmics a través d'una plataforma pública sembla indispensable per tal que es pugui fer front a la discriminació algorítmica, i més específicament, dur a terme de manera efectiva l'acció col·lectiva en matèria de no-discriminació algorítmica.

Hodges es pregunta, en l'àmbit del dret de les persones consumidores, com s'identifiquen els problemes massius, i conclou que usualment és a través de la presentació, davant un organisme independent, de queixes per part de persones consumidores individuals. Llavors, aquest organisme independent examina les queixes i identifica que n'hi ha de similars, establint la concurrència d'un problema sistèmic¹⁷⁹⁶. En aquest sentit, referencia el Belgian Consumer Mediation Service, que en l'àmbit del dret de les persones consumidores, té un únic portal o punt d'accés que dona consell i accés a sistemes de protecció i que ha demostrat poder identificar problemes sistèmics ràpidament perquè atrau un nombre suficient de contactes individuals, pels quals el problema en qüestió pot ser identificat via web i ràpidament¹⁷⁹⁷. Per tant, el fet que existeixi una única plataforma que centralitzi la informació sobre queixes de persones en relació amb la discriminació algorítmica dins d'un mateix estat membre pot possibilitar la detecció de problemes sistèmics. El fet que la gestioni l'organisme d'igualtat també possibilita que amb la informació obtinguda pugui exercir les seves funcions, com també notificar-ho a altres autoritats nacionals o europees¹⁷⁹⁸ -com aquelles que s'encarreguen del control sobre la protecció de dades o dels sistemes d'intel·ligència artificial. El mateix autor fa referència també a la figura de l'Ombudsperson, que pot identificar l'existència de demandes individuals similars i agregar-les per donar decisions coherents (també identificar problemes sistèmics), o desenvolupar en sectors específics casos de prova en els tribunals per veure l'èxit potencial de casos similars¹⁷⁹⁹.

En aquest sentit, una plataforma telemàtica pública gestionada per l'organisme d'igualtat podria facilitar la detecció de problemes sistèmics, la interrelació de persones que s'han vist discriminades mitjançant sistemes algorítmics i possibilitar l'inici de litigis en favor de les persones que han emprat la plataforma pública per cessar sistemes algorítmics amb

¹⁷⁹⁶ HODGES, C., "Collective Redress: The Need for New Technologies", *op. cit.*, 2019, pàg. 84.

¹⁷⁹⁷ *Ídem*, pàg. 84.

¹⁷⁹⁸ *Ídem*, pàg. 68. L'autor especifica que és una pràctica estàndard al Regne Unit en molts sectors que la detecció d'un problema sistèmic generi la coordinació entre qui ha detectat el problema i altres autoritats nacionals que també han de col·laborar per erradicar-lo.

¹⁷⁹⁹ *Ídem, ibídem*.

resultats discriminatoris.

2.2.4. Litigació: Acció de cessació de sistemes algorítmics que tinguin resultats discriminatoris

Tal com hem esmentat anteriorment, en l'article 7.2 de la Directiva 2000/43/CE, 9.2 de la Directiva 2000/78/CE, l'article 8.3 de la Directiva 2004/113/CE i 17.2 de la Directiva 2006/54/CE s'estableix que els estats membres vetllaran per tal que les associacions, organitzacions i altres persones jurídiques (...) puguin iniciar, en nom de la persona demandant o en el seu suport i la seva autorització, qualsevol procediment judicial o administratiu previst per tal d'exigir el compliment de les obligacions de la Directiva¹⁸⁰⁰.

La Recomanació 2018/951 estableix que els estats membres haurien de considerar que els organismes d'igualtat representessin a les persones perjudicades per discriminació davant dels tribunals¹⁸⁰¹, en el seu nom o en el de les víctimes, podent dur a terme litigació estratègica¹⁸⁰² i donant-los poders per obtenir proves i informació rellevant que pot donar suport a la part en el procés legal¹⁸⁰³. També l'organisme Comissió Europea contra el Racisme i la Intolerància (d'ara en endavant, ECRI), va adoptar la "Recomendación núm. 2 de política general de la ECRI: organismos de promoción de la igualdad para luchar contra el racismo y la intolerancia a nivel nacional", el 7 de desembre de 2017¹⁸⁰⁴ (que revisava la seva Recomanació núm. 2 de política general adoptada l'any 1997¹⁸⁰⁵), on establí que s'hauria d'assignar als organismes de promoció de la igualtat la funció de litigar en nom de les persones perjudicades per discriminació¹⁸⁰⁶, dur a terme litigació

¹⁸⁰⁰ Article 7.2 de la Directiva 2000/43/CE; article 9.2 de la Directiva 2000/78/CE; article 8.3 de la Directiva 2004/113/CE; i article 17.2 de la Directiva 2006/54/CE.

¹⁸⁰¹ Apartat 1.1.2 (1) de la Recomendación (UE) 2018/951.

¹⁸⁰² Apartat 1.1.2 (2) de la Recomendación (UE) 2018/951.

¹⁸⁰³ Apartat 1.1.2 (4) de la Recomendación (UE) 2018/951.

¹⁸⁰⁴ COMISIÓN EUROPEA CONTRA EL RACISMO Y LA INTOLERANCIA (ECRI), *Recomendación núm. 2 de política general de la ECRI: organismos de promoción de la igualdad para luchar contra el racismo y la intolerancia a nivel nacional*, Adoptada el 7 de diciembre de 2017, CRI(2018)06, Estrasburgo, 27 de febrero de 2018.

¹⁸⁰⁵ COMISIÓN EUROPEA CONTRA EL RACISMO Y LA INTOLERANCIA (ECRI), *Recomendación núm. 2 de política general de la ECRI: los órganos especializados en la lucha contra el racismo, la xenofobia, el antisemitismo y la intolerancia a nivel nacional*, Estrasburgo, 13 de junio de 1997.

¹⁸⁰⁶ Apartat 14 de la Recomendación núm. 2 de política general de la ECRI (2017). En el sentit d'instar als estats membres de la UE a facilitar als organismes d'igualtat la legitimació en casos de discriminació davant

estratègica¹⁸⁰⁷, i obtenir proves i informació¹⁸⁰⁸.

Kádár defensa el potencial de la legitimació dels organismes d'igualtat per litigar en casos de discriminació, tant per millorar l'aplicació de la normativa antidiscriminació com l'accés a la justícia per part de les víctimes, podent-se dur a terme la representació davant el tribunal per part del propi personal de l'organisme d'igualtat¹⁸⁰⁹. Allen i Masters també proposen que els organismes d'igualtat haurien de considerar dur a terme litigació estratègica per desafiar els sistemes d'IA discriminatoris, tant per donar suport a persones individuals, com per evidenciar que l'exercici regulador i el principi d'igualtat i no-discriminació són efectius¹⁸¹⁰, i Farkas proposa que el model ideal per assegurar l'aplicació de la normativa antidiscriminació es focalitzaria en agències i organismes públics nacionals i europeus especialitzats, amb poders per investigar i llançar accions de representació grupals o col·lectives que busquin canvis sistèmics¹⁸¹¹. La legitimació dels organismes d'igualtat en casos de discriminació hauria d'incloure representar víctimes de discriminació davant dels tribunals (amb consentiment o aprovació per part de la víctima), iniciar procediments en el seu nom (quan no existeix una víctima identificable, incloent l'acció popular, o la representació dels interessos d'una víctima que no pren part en el procés judicial, amb el seu consentiment o aprovació), intervenir per donar suport a una part, participar com a *amicus curiae* o defensar les seves decisions legalment vinculants davant els tribunals¹⁸¹².

Ja existeixen organismes nacionals amb llarg recorregut que tenen entre les seves funcions iniciar accions col·lectives. Hodges fa referència a diverses autoritats públiques reguladores de certs estats membres de la UE que poden dur a terme accions de litigació col·lectiva, com el “Forbrugerombudsmanden” danès, anomenat usualment “Consumer Ombudsman” en anglès, que pot portar una acció col·lectiva al Tribunal de Mercats en nom de consumidors afectats en un sistema “opt-out” per reclamacions fins a 270 euros i

els tribunals, vegeu també: COUNCIL OF EUROPE PARLIAMENTARY ASSEMBLY, *Resolution 2054 (2015) Equality and non-discrimination in the access to justice*, Strasbourg, 24 April 2015, punt 5.4; COMMISSIONER FOR HUMAN RIGHTS, *Opinion of the Commissioner for Human Rights on national structures for promoting equality*, CommDH(2011)2, Council of Europe, Strasbourg, 21 March 2011, pàg. 16.

¹⁸⁰⁷ Apartat 15 de la *Recomendación núm. 2 de política general de la ECRI (2017)*.

¹⁸⁰⁸ Apartat 21 de la *Recomendación núm. 2 de política general de la ECRI (2017)*.

¹⁸⁰⁹ KÁDÁR, T., “The legal standing of equality bodies”, *op. cit.*, 2019, pàg. 1.

¹⁸¹⁰ ALLEN, R. i MASTERS, D., *Regulating for an equal AI...*, *op. cit.*, 2020, pàg. 23.

¹⁸¹¹ FARKAS, L., “Collective actions under European anti-discrimination law”, *op. cit.*, 2014, pàg. 39.

¹⁸¹² KÁDÁR, T., “The legal standing of equality bodies”, *op. cit.*, 2019, pàg. 1 i 2.

més enllà d'aquesta suma, en un sistema “opt-in”. A Finlàndia, una autoritat equivalent només té el poder de dur a terme accions col·lectives del sistema “opt-in” i ha emprat aquest poder per negociar pagaments de reparació¹⁸¹³. L'autor, en l'àmbit del consum, i extrapolable a la discriminació algorítmica, ha fet una anàlisi exhaustiva del concepte de compensació reguladora¹⁸¹⁴, que suposa en l'esquema regulador la participació d'un organisme públic que acusi a una empresa o organització d'incompliment normatiu (en l'àmbit del consum, per exemple, infraccions en la publicitat, subministrament d'informació, seguretat del producte, etc.; i en l'àmbit de la discriminació algorítmica, per exemple, incompliment de la normativa antidiscriminació). Arran d'aquest incompliment normatiu, l'organisme pot, o bé requerir el cessament de l'acció, o bé requerir que l'empresa o organització dugui a terme accions per evitar que es repeteixi l'acció (o per tal que es redueixi el risc), o bé obligar al pagament de reparació a les persones perjudicades (posa com a exemple el British Financial), o bé procedir a iniciar una acció de representació (Consumer Ombudsman Danès)¹⁸¹⁵. Conclou que l'èxit d'aquestes autoritats resideix parcialment en què no només disposen de la possibilitat de litigar, sinó que també disposen de múltiples eines (entre les quals s'inclouen típicament ordenar el cessament d'una conducta, investigar i obtenir evidències, requerir el canvi d'un comportament, requerir compensació i/o imposar sancions) per iniciar un procediment envers una empresa que ha incomplert la normativa¹⁸¹⁶.

En el sentit de tenir la facultat d'ordenar el cessament d'una conducta, i com a funció d'interès per tenir eines enfront la discriminació algorítmica, l'article 8.4 de la Directiva 2020/1828 faculta els estats membres a introduir com a requisit d'admissibilitat per part del tribunal de l'acció de cessació, la consulta prèvia de l'entitat habilitada a la persona empresària a mode de requeriment per a que cessi la conducta infractora abans d'interposar una acció judicial. Una de les funcions dels organismes d'igualtat encarregats de dur a terme accions col·lectives de cessació de conductes discriminatòries mitjançant sistemes algorítmics podria ser requerir l'empresa, organització o administració pública prèviament a l'acció judicial, que cessés la discriminació algorítmica en base a certa

¹⁸¹³ HODGES, C., “Collective Redress: The Need for New Technologies”, *op. cit.*, 2019, pàg. 63.

¹⁸¹⁴ En anglès, “regulatory redress”.

¹⁸¹⁵ HODGES, C., “Collective Redress: The Need for New Technologies”, *op. cit.*, 2019, pàg. 63.

¹⁸¹⁶ *Ídem*, pàg. 63.

categoria protegida o certes categories protegides i aportés proves dels canvis que ha dut a terme.

En relació amb el fet que un mateix organisme d'igualtat pugui dur a terme tant una funció de litigació com quasi-jurisdiccional, és a dir, per una banda, actuar en favor d'una part en un litigi, i per l'altra actuar amb imparcialitat per prendre una decisió pròpia com a organisme quasi-jurisdiccional¹⁸¹⁷, s'especifica en la Recomanació núm. 2 de política general de la ECRI: Organismes de promoció de la Igualtat per lluitar contra el racisme i la intolerància a escala nacional, de 7 de desembre de 2017, que és necessari assegurar que cada funció la dugui a terme una unitat diferent amb personal diferent i recursos financers i humans apropiats¹⁸¹⁸. Crowley considera que la combinació en un únic organisme d'igualtat de la funció de litigació i presa de decisions (com també promoció i prevenció de la discriminació i suport a les víctimes), pot suposar l'existència de tensions dins l'organisme per dur-les a terme de manera adequada si hi ha recursos limitats¹⁸¹⁹. Tanmateix, Kádár constata que, a la pràctica, aquells organismes d'igualtat que realitzen tant la competència de litigació com d'actuar com a òrgan quasi-jurisdiccional tenen una visió clara sobre com prioritzar una o l'altra i construeixen salvaguardes entre les unitats que duen a terme les diferents competències¹⁸²⁰. Sigui quina sigui la percepció sobre la seva adequació, hi ha organismes d'igualtat europeus que duen a terme aquestes funcions alhora, tal com l'Unia (Interfederal Centre for Equal Opportunities) a Bèlgica, el National Council for Combating Discrimination a Romania o l'Equality Ombudsman a Suècia¹⁸²¹.

En resum, una funció de l'organisme d'igualtat que pot tenir especial incidència en la lluita envers la discriminació algorítmica és la presentació d'accions de cessació amb l'objectiu de fer cessar l'ús de sistemes algorítmics discriminatoris. Aquesta acció es podria dur a terme per aconseguir que no es pogués usar -tant en el moment com en el futur- el sistema algorítmic discriminatori (posteriorment, les persones perjudicades pel sistema algorítmic discriminatori podrien demanar la reparació pel perjudici patit). Tant la posició de privilegi respecte a les persones individualment considerades per obtenir informació de rellevància en el camp de la discriminació algorítmica, com la possibilitat

¹⁸¹⁷ Apartat 45 del Memorándum explicativo de la *Recomendación núm. 2 de política general de la ECRI* (2017).

¹⁸¹⁸ Apartat 11 de la *Recomendación núm. 2 de política general de la ECRI* (2017).

¹⁸¹⁹ CROWLEY, N., *Equality bodies making a difference*, op. cit., 2018, pàg. 47.

¹⁸²⁰ KÁDÁR, T., "The legal standing of equality bodies", op. cit., 2019, pàg. 14.

¹⁸²¹ *Ídem*, pàg. 7-14. L'autor fa una anàlisi tant de les funcions dels tres organismes referenciats com del seu impacte com a organismes de referència en la lluita envers la discriminació.

de gestionar i confeccionar estadístiques que puguin servir com a elements comparatius en cas de discriminació indirecta algorítmica, com la possibilitat de gestionar una plataforma pública que faciliti la gestió de queixes per discriminació algorítmica, interrelacioni persones perjudicades i pugui detectar problemàtiques sistèmiques per iniciar litigació estratègica per part dels organismes d'igualtat (que en molts estats membres de la UE estan legitimats per dur a terme accions judicials i administratives en matèria antidiscriminació), són elements que fan pensar que els organismes d'igualtat dels estats membres són els més ben posicionats per ser els actors principals del sistema col·lectiu-institucional de lluita envers la discriminació algorítmica.

2.3. Altres funcions

Hi ha altres funcions que l'organisme d'igualtat pot dur a terme per millorar la lluita envers la discriminació algorítmica. Estudiem en primer lloc l'abordatge de la discriminació algorítmica interseccional i múltiple (2.3.1), la participació en els espais de prova dels sistemes d'intel·ligència artificial que es vulguin incorporar en el mercat intern europeu (2.3.2) i finalment, la gestió d'una llista pública de sistemes algorítmics discriminatoris (2.3.3).

2.3.1. Abordatge unificat de la discriminació algorítmica interseccional

Una de les temàtiques que hem tractat recurrentment en aquest treball ha estat la rellevància de la interseccionalitat en la discriminació algorítmica i la discriminació múltiple (vegeu apartat 2.2 del capítol IV). Tal com hem analitzat, l'ús de sistemes algorítmics possibilita que aquesta es doni de manera predominant atesa la gran capacitat d'interrelacionar diferents dades amb la qual compten els sistemes algorítmics. Que els organismes d'igualtat dels estats membres no estiguin dividits segons la característica protegida, sinó que abastin en un mateix organisme la protecció de totes les característiques protegides per la legislació, possibilita que tota la

informació que obtinguin per part de persones que s'hagin vist discriminades pugui ser interrelacionada i llegida en clau interseccional. Tal com estableixen Chopin i Germaine, la instauració d'organismes que abasten molts atributs protegits no només fa més efectiva la gestió de casos interseccionals i de discriminació múltiple, sinó que també suposa un avantatge en l'efectivitat del cost-benefici i facilita l'accés a persones que s'han vist discriminades per fer queixes (per exemple, la visibilitat i identificació d'un únic organisme d'igualtat constitueix un accés a la protecció més senzill que no si existissin diversos organismes segons l'atribut protegit)¹⁸²².

La interrelació de la informació obtinguda per l'organisme d'igualtat en clau interseccional també ha d'implicar la informació estadística amb la qual treballi l'organisme. Tal com hem plantejat en l'apartat on ens referíem a la confecció d'estadístiques per part dels organismes d'igualtat, aquestes poden ser de molta rellevància tant per poder evidenciar la situació en referència a la discriminació en els estats membres, com per ser emprades com a elements comparatius en casos de discriminació algorítmica indirecta. Doncs bé, per assegurar la comprensió de les dades d'igualtat s'ha de tenir en compte la discriminació múltiple i interseccional en les dades i la seva anàlisi. Per exemple, dades recollides en l'àrea de la discapacitat han d'incloure informació sobre gènere, raça o origen ètnic, edat, o altres característiques potencials que puguin incrementar el risc de discriminació basada en múltiples factors (aquest tipus d'anàlisi generalment requereix bases de dades més àmplies per una anàlisi robusta, diferents fonts de dades que puguin ser considerades en grans enquestes, i dades administratives i vinculades)¹⁸²³.

En definitiva, com ja hem abordat en l'apartat 4.3.2 del capítol II, el fet que la discriminació algorítmica faciliti la creació de grups de persones discriminades més atomitzats, atès que un grup creat per una combinació de característiques protegides serà més petit i més difícilment identificable que un grup creat únicament per una característica protegida (comparem, per exemple, que el sistema algorítmic discrimini a dones, gitanes de més de 65 anys amb que discrimini a totes les dones), dificulta tant la

¹⁸²² EUROPEAN COMMISSION, EUROPEAN COMMISSION, *Equality bodies and the implementation...*, *op. cit.*, 19 May 2021, pàg. 5; CHOPIN, I. i GERMAINE, C., *A comparative analysis of...*, *op. cit.*, 2022, pàg. 108.

¹⁸²³ HIGH LEVEL GROUP ON NON-DISCRIMINATION, EQUALITY, AND DIVERSITY. SUBGROUP ON EQUALITY DATA, *Guidelines on improving the collection and use of equality data*, Brussels, 2018, pàg. 18.

detecció de la discriminació com l'existència d'elements de comparació. Per consegüent, la tasca de l'organisme d'igualtat pot suposar indagar en les tipologies de discriminacions múltiples i interseccionals, investigar a través d'estadístiques la discriminació causada arran de la combinació d'atributs i els efectes que pot tenir en la societat (que segurament opera de manera encara més subtil que si només està implicada una única característica protegida), investigar aquestes situacions gràcies a l'obtenció de dades centralitzades (per exemple, interrelacionar persones que poden patir aquest tipus de discriminació i operar com a entitat coordinadora per la defensa dels seus interessos) i inclús provar els sistemes algorítmics, com veurem en el següent apartat, per tal de verificar que si bé aparentment el sistema algorítmic no és discriminatori en base a una característica protegida, tampoc ho és en base a la combinació de diferents característiques protegides.

2.3.2. Participar en els espais de prova dels sistemes algorítmics

Tal com hem establert en l'apartat 3.5 del capítol VI del treball, la PRIA estableix en el seu article 53.1 la institució d'espais controlats de prova per a la IA per proporcionar un entorn controlat que faciliti el desenvolupament, la prova i la validació de sistemes innovadors d'IA¹⁸²⁴. En el segon apartat de l'article, s'estableix que els estats membres vetllaran per tal que en la mesura en què els sistemes innovadors d'IA impliquin el tractament de dades personals o estiguin inclosos en l'àmbit de supervisió d'altres autoritats nacionals, aquestes últimes hauran d'estar lligades al funcionament de l'espai controlat de proves de la IA¹⁸²⁵. En aquest sentit, una de les tasques de l'organisme d'igualtat ha de ser prendre part en aquests espais de prova i vetllar per tal que quan s'estigui experimentant amb un sistema d'IA, aquest no presenti resultats discriminatoris en relació amb grups determinats de població o inclús relacionant diferents característiques protegides.

Aquests espais de prova suposen que el desenvolupament dels sistemes d'IA es faci sota una vigilància reguladora abans de la seva introducció en el mercat o posada en servei¹⁸²⁶.

¹⁸²⁴ Article 53.1 de la PRIA.

¹⁸²⁵ Article 53.2 de la PRIA.

¹⁸²⁶ Considerant 71 de la PRIA.

En el cas de la discriminació algorítmica, com ja hem apuntat, aquesta es pot desenvolupar també quan el sistema algorítmic s'aplica a diferents contextos. Per tant, que el sistema no mostri signes de proporcionar resultats discriminatoris en un espai controlat de prova no obsta que posteriorment, quan s'apliqui en un context determinat, sí que acabi mostrant resultats discriminatoris. Tanmateix, el fet que en els espais de prova es detecti la possibilitat de discriminació i s'obligui l'empresa proveïdora del sistema a mitigar el risc de discriminació pot ser una eina molt valuosa per lluitar envers la discriminació algorítmica, ja que es podrà verificar que els sistemes algorítmics no són discriminatoris abans d'emprar-los. Anant més enllà, per assegurar que tots els sistemes parteixen d'una base no discriminatòria, s'hauria d'obligar a qualsevol sistema d'intel·ligència artificial que es volgués aplicar a la presa de decisions en el context social i que es volgués introduir en el mercat a passar per l'espai de prova, i així les autoritats de vigilància competents, ja sigui l'AINSA, l'AEPD o l'Autoritat Independent, poguessin verificar que cap sistema algorítmic nou introduït en el mercat provoca discriminació algorítmica. Actualment la regulació, com hem vist, és molt més permissiva.

Una altra funció de l'organisme d'igualtat de rellevància en relació amb la discriminació algorítmica podria ser la gestió d'una llista pública dels sistemes algorítmics discriminatoris, com veurem tot seguit.

2.3.3. Llista pública dels sistemes algorítmics discriminatoris

La gestió d'una llista dels sistemes algorítmics que s'ha establert ja sigui per resolució judicial o per decisió de l'organisme d'igualtat que són discriminatoris esdevé també un element clau per la lluita envers la discriminació algorítmica.

Aquesta llista hauria d'estar actualitzada i ser accessible al públic de la mateixa manera que la base de dades de la UE per a sistemes d'IA de risc alt independents¹⁸²⁷. Per tal que aquesta llista sigui operativa s'han de donar dues condicions prèvies (vegeu apartat 4 del capítol VII): en primer lloc, interessa que s'obligui els proveïdors de sistemes d'IA a que anomenin el seu sistema algorítmic de manera que es pugui reconèixer per part de les

¹⁸²⁷ Article 60 de la PRIA.

persones sobre les quals ha pres decisions; i en segon lloc, i establerta una base de dades sobre els sistemes algorítmics de presa de decisions en contextos socials existents -no només aquells de risc alt, sinó tots els que operin en el mercat-, que s'obligui a les empreses, organitzacions i administracions públiques a publicar un catàleg dels sistemes algorítmics que estan utilitzant i una explicació de per què els estan utilitzant (per exemple, en una pàgina web accessible al públic general que s'especifiqui que s'està emprant el sistema algorítmic "x" per gestionar els processos selectius de l'empresa)¹⁸²⁸.

La llista pública sobre els sistemes algorítmics discriminatoris no només pot ajudar les persones que s'han vist discriminades a localitzar si el sistema algorítmic amb el qual s'ha pres una decisió sobre elles s'ha considerat discriminatori (entenem en aquest cas que la persona coneix quin sistema algorítmic s'ha utilitzat per prendre una decisió sobre ella i que un mateix sistema algorítmic pot ser utilitzat per diferents empreses i organitzacions), sinó que també genera transparència i accés a informació per part del públic general. En aquest sentit, interessa referenciar que l'Observatori d'Algoritmes amb Impacte Social (OASI), desenvolupat per Eticas Foundation, ha creat un registre públic amb informació sobre certs algoritmes (com s'anomena, descripció, països on s'aplica, en quin àmbit, quin impacte social té, etc.) per incrementar la conscienciació pública¹⁸²⁹.

La tasca de mantenir una llista només la pot dur a terme un organisme que tingui informació centralitzada i actualitzada sobre les resolucions dels casos de discriminació i

¹⁸²⁸ Vegeu: Decreto-ley 2/2023, de 8 de marzo, de medidas urgentes de impulso a la inteligencia artificial en Extremadura. DOE, núm. 48, de 10 de marzo de 2023. En aquest decret llei s'estableix que quan s'adoptin actes administratius mitjançant sistemes d'intel·ligència artificial en el marc d'un procediment administratiu, a més dels requisits prevists en l'article 41 de la Ley 40/2015, de 1 de octubre, de Régimen Jurídico del Sector Público, es donarà la deguda publicitat del mecanisme de decisió, de les prioritats assignades en el procediment d'avaluació i de la presa de decisions, així com totes les dades que puguin impactar en el seu contingut (art. 11.1 i 2). També s'estableix que les normes que regulin els procediments administratius faran referència expressa a l'impacte de l'ús d'aquests sistemes en la prestació de serveis públics que, en el seu cas, suportin la presentació de sol·licituds, declaracions responsables o comunicacions, la comprovació o verificació dels requisits de les persones interessades, així com la presa de decisions (art. 12.1). Es deixarà també constància en l'Inventari d'Informació Administrativa, sobre les validacions realitzades per l'òrgan responsable dels procediments administratius respecte el procés de lògic dissenyat per la realització dels actes, els riscos que implica i qualssevol altres aspectes que garanteixin els drets de les persones interessades (art. 12.2). Vegeu també: Llei 1/2022, de 13 d'abril, de la Generalitat, de Transparència i Bon Govern de la Comunitat Valenciana. DOGV, núm. 9323, de 22 d'abril de 2022. En el seu article 16.1.1 s'estableix que les administracions públiques hauran de publicar "*la relació de sistemes algorítmics o d'intel·ligència artificial que tinguen impacte en els procediments administratius o la prestació dels serveis públics amb la descripció de manera comprensible del seu disseny i funcionament, el nivell de risc que impliquen i el punt de contacte al qual poder adreçar-se en cada cas, d'acord amb els principis de transparència i explicabilitat*".

¹⁸²⁹ Per visitar el registre, vegeu: ETICAS FOUNDATION, Observatory of Algorithms with Social Impact. Disponible a: <https://eticas.tech/es/oasi>. Darrera visita el 2 de setembre de 2023.

que proactivament promogui que s'examinin i es supervisin els sistemes algorítmics -a través d'investigacions pròpies i accions de cessació, per exemple-, per tal d'assegurar que els resultats que generen són el mínim discriminatoris possible.

Recapitulant, l'organisme nacional d'igualtat com a òrgan quasi-jurisdiccional per lluitar envers la discriminació algorítmica ha de tenir la possibilitat d'iniciar investigacions pròpies, prendre decisions vinculants i imposar sancions si hi ha violacions del principi de no-discriminació. Com a organisme amb possibilitat de litigar en favor de persones i col·lectius que pateixen discriminació algorítmica, ha de poder obtenir informació de rellevància en el camp de la discriminació algorítmica, gestionar i confeccionar estadístiques que puguin servir com a elements comparatius en cas de discriminació indirecta algorítmica, gestionar una plataforma pública que faciliti la gestió de queixes per discriminació algorítmica, interrelacionar persones perjudicades i detectar problemàtiques sistèmiques per iniciar litigació estratègica i iniciar accions de cessació amb l'objectiu de fer cessar l'ús de sistemes algorítmics discriminatoris. Poden posicionar-se de manera privilegiada, addicionalment, per abordar de manera unificada la discriminació interseccional, participar en els espais de prova dels sistemes algorítmics que preveu la PRIA i mantenir una llista de sistemes algorítmics que s'hagin provat discriminatoris. A continuació, estudiem, en el cas espanyol, com s'ha configurat l'organisme d'igualtat, si s'adapta als estàndards europeus, i quines competències i poders té per fer front a la discriminació algorítmica.

3. El cas espanyol: Autoritat independent per a la Igualtat de Tracte i la No Discriminació en el context de la discriminació algorítmica

En aquest apartat volem analitzar si l'Autoritat Independent, creada per la Llei 15/2022, que actuarà com a organisme d'igualtat espanyol, s'adapta a la configuració que planteja la UE pels organismes d'igualtat (3.1), i quines de les funcions que se li han atorgat són útils per fer front a la discriminació algorítmica (3.2), segons l'anàlisi que hem fet en el segon apartat d'aquest capítol.

3.1. L'adaptació de l'organisme d'igualtat espanyol als estàndards europeus: l'Autoritat Independent per a la Igualtat de Tracte i la No Discriminació

L'article 33 de la Llei 62/2003 transposa a l'ordenament jurídic espanyol les exigències establertes en l'article 13 de la Directiva 2000/43/CE, i crea el Consejo para la Eliminación de la Discriminación Racial o Étnica (d'ara en endavant, CEDRE)¹⁸³⁰. Aquest és l'òrgan precursor de l'Autoritat Independent, creada per la Llei 15/2022 i que pretén adaptar-se millor a les exigències europees en tant que organisme d'igualtat. El CEDRE no s'ajusta als perfils clars que exigeix la UE en matèria d'organismes d'igualtat¹⁸³¹, per raons d'escassetat i precarietat de recursos, escàs coneixement públic¹⁸³², falta de coneixement per part de les víctimes potencials de la seva existència, impossibilitat de posar sancions, absència de legitimació per litigar en nom -ni tampoc al marge- de les víctimes, i falta d'independència i autonomia¹⁸³³ (en el cinquè informe de l'ECRI de seguiment sobre Espanya s'afirma que no té un organisme independent de promoció per a la igualtat capaç d'introduir un canvi social real i d'aconseguir una major igualtat, i recomana que se'n creï un¹⁸³⁴, donant per fet que el CEDRE no opera com a tal¹⁸³⁵). Tampoc s'adapta als estàndards europeus ja que només protegeix la discriminació racial i no els altres atributs protegits, a diferència de la majoria dels seus homòlegs europeus¹⁸³⁶.

¹⁸³⁰ Article 33 de la Llei 62/2003.

¹⁸³¹ REY MARTÍNEZ, F., "O organismo español para a igualdad...", *op. cit.*, 2021, pàg. 314. L'autor opina que aquesta inadaptació es deu a que a Espanya no s'ha percebut la discriminació racial, en general, com a un problema i referencia tres sentències del TC on el tribunal es mostra totalment indiferent al factor racial. Vegeu: Tribunal Constitucional (Sala Segunda), sentència 13/2001, de 29 de enero de 2001; Tribunal Constitucional (Sala Primera), sentència 69/2007, de 16 de abril de 2007 i; Tribunal Constitucional (Sala Segunda), sentència 1/2021, de 25 de enero de 2021.

¹⁸³² SEBASTIANI, L., "Investigando los límites de la lucha...", *op. cit.*, 2021, pàg. 770.

¹⁸³³ REY MARTÍNEZ, F., "O organismo español para a igualdad...", *op. cit.*, 2021, pàg. 318-321.

¹⁸³⁴ COMISIÓN EUROPEA CONTRA EL RACISMO Y LA INTOLERANCIA (ECRI), *Informe de la ECRI sobre España. Quinto ciclo de supervisión*, Adoptado el 5 de diciembre de 2017, CRI(2018)2, Estrasburgo, 27 de febrero de 2018, pàg. 18-20. El cinquè informe de seguiment de l'ECRI estableix que a Espanya hi manca un organisme independent de promoció capaç d'introduir un canvi social real i d'aconseguir una major igualtat. El CEDRE no és independent, no té infraestructura, ni ha realitzat cap activitat important en els últims anys, ni té facultat d'investigar casos, ni pot iniciar o acompanyar víctimes en procediments judicials. En l'informe s'insta al CEDRE a reformar els seus estatuts per poder assegurar la seva independència del govern i poder complir amb les competències que li assigna el model europeu.

¹⁸³⁵ REY MARTÍNEZ, F., "O organismo español para a igualdad...", *op. cit.*, 2021, pàg. 322.

¹⁸³⁶ REY MARTÍNEZ, F., *Derecho Antidiscriminatorio*, *op. cit.*, 2019, pàg. 131. A Catalunya, cal fer esment a l'Organisme de protecció i promoció de la igualtat de tracte i la no discriminació, creat per la Llei 19/2020, que és un òrgan transversal que protegeix totes les característiques protegides sota la llei; REY MARTÍNEZ, F., "O organismo español para a igualdad...", *op. cit.*, 2021, pàg. 325, 326 i 335. L'autor planteja que, malgrat la norma sigui un organisme que protegeix tots els atributs protegits, existeix debat sobre si la igualtat de gènere s'ha de protegir amb un organisme d'igualtat diferent al general, atès que entre altres coses, les dones no són una minoria social, sinó al contrari, una majoria.

La Llei 15/2022 reconeix a l'exposició de motius que la transposició de la Directiva 2000/43/CE anterior era insuficient i ineficient¹⁸³⁷. Per posar-hi remei, entre d'altres disposicions, crea l'Autoritat Independent, encarregada de promoure la igualtat de tracte i la no discriminació tant en el sector públic com en el sector privat¹⁸³⁸, amb múltiples tasques i que inclou tots els motius de discriminació establerts per la llei, una perspectiva que, com ja hem esmentat, permet gestionar millor els casos de discriminació múltiple o interseccionals¹⁸³⁹. Es configura com una entitat de dret públic dotada de personalitat jurídica pròpia i plena capacitat pública i privada, que actua per al compliment de les seves finalitats amb plena independència i autonomia funcional respecte les administracions públiques¹⁸⁴⁰. Això significa que l'Autoritat Independent podrà dirigir-se tant contra un sistema algorítmic discriminatori emprat per una empresa privada com una administració pública. La independència de l'Autoritat Independent de l'Administració Pública és d'especial rellevància per tal que pugui actuar tant en relació amb sistemes algorítmics que s'emprin en el sector públic com en el sector privat.

Es regirà pel seu propi Estatut (que contindrà l'estructura orgànica depenent de l'Autoritat Independent, el seu règim de funcionament intern, el seu règim de personal, el seu règim econòmic i pressupostari i totes les altres qüestions relatives al seu funcionament i règim d'actuació¹⁸⁴¹), i les normes de dret administratiu general i especial que li siguin d'aplicació, entenent-se llavors que les seves resolucions posen fi a la via administrativa¹⁸⁴². Dirigida i representada per la persona que ocupi la seva presidència, que serà nomenada pel Govern mitjançant Reial Decret, el seu mandat serà per cinc anys sense possibilitat de renovació¹⁸⁴³. També disposa la llei que l'Autoritat Independent cooperarà en l'exercici de les seves competències amb organismes públics (com podria ser l'AEPD o l'AESIA)¹⁸⁴⁴ i la Defensoria del Poble (sense perjudici de les competències

¹⁸³⁷ Exposició de motius de la Llei 15/2022.

¹⁸³⁸ Article 40 de la Llei 15/2022.

¹⁸³⁹ REY MARTÍNEZ, F., "O organismo español para a igualdad...", *op. cit.*, 2021, pàg. 335.

¹⁸⁴⁰ Article 41.1 de la Llei 15/2022.

¹⁸⁴¹ Article 41.3 de la Llei 15/2022. L'article 43 de la mateixa norma estableix que l'Estatut regularà les formes i el procediment per assegurar la participació en les seves activitats de les organitzacions representatives i dels interessos socials afectats.

¹⁸⁴² Article 41.2 de la Llei 15/2022.

¹⁸⁴³ Article 41.4 de la Llei 15/2022. L'article 42.1 de la mateixa norma també estableix que el personal al servei de l'Autoritat independent serà amb caràcter general funcionari de carrera de les administracions públiques o personal laboral procedent d'organismes nacionals o internacionals amb funcions en matèria d'igualtat de tracte i no-discriminació i que en els processos selectius s'incorporaran mesures d'acció positiva en benefici de les persones protegides per la llei.

¹⁸⁴⁴ Article 44.2 de la Llei 15/2022.

d'aquest últim òrgan)¹⁸⁴⁵, i s'establirà la col·laboració entre administracions públiques¹⁸⁴⁶.

Rey Martínez observa que els trets fonamentals que presenta l'Autoritat Independent (autonomia organitzativa, pressupostària, desenvolupament normatiu, nomenament governamental -però amb prèvia aprovació parlamentària¹⁸⁴⁷-, i causes de cessament taxades), sí que semblen assegurar la seva independència respecte del govern i responen al marc de la Recomanació 2018/951. Tanmateix, tot i que l'Autoritat Independent compleix amb els estàndards europeus, ho fa d'un mode bastant modest, és a dir, de manera escassa respecte a les seves possibilitats de configuració¹⁸⁴⁸.

Estudiem, a continuació, les funcions atorgades a l'Autoritat Independent i l'ús d'aquestes en casos de discriminació algorítmica.

3.2. Funcions de l'Autoritat Independent amb incidència en matèria de discriminació algorítmica

¹⁸⁴⁵ Article 53 de la Llei 15/2022.

¹⁸⁴⁶ Article 35 de la Llei 15/2022. La llei 19/2020 també estableix en el seu article 29 la col·laboració entre institucions públiques.

¹⁸⁴⁷ Malgrat que no s'arriba a l'ideal de nomenament per concurs, s'exigeix que la persona que dirigeixi l'autoritat sigui de reconegut prestigi en defensa i promoció de la igualtat.

¹⁸⁴⁸ REY MARTÍNEZ, F., "O organismo español para a igualdad...", *op. cit.*, 2021, pàg. 331 i 339. Segons l'autor, la millor manera com s'hauria d'haver configurat l'organisme d'igualtat espanyol és a través de la figura de l'Ombudsperson (és a dir, la defensoria del poble), ateses les seves avantatges com les sinergies entre funcions, aprenentatge mutu, estalviar costos, ampliar l'equip humà qualificat, major disponibilitat per a les víctimes i més facilitat de ser coneguts per el públic. També al complir perfectament el requisits d'independència que exigeix el dret comunitari (atès que el nomenament parlamentari és amb majoria reforçada) i amb capacitat d'assolir més visibilitat ciutadana que l'Autoritat Independent. També allà on existissin defensories autonòmiques (per exemple, a Catalunya, la Sindicatura de Greuges), aquestes podrien assumir, previ conveni amb la defensoria estatal, les funcions dels organismes per a la igualtat en els àmbits autonòmics i local, reservant la defensoria estatal a la defensa de la igualtat enfront l'administració estatal i institucional. Referència l'autor l'organisme d'igualtat a França, on es va configurar inicialment com una institució especialitzada, l'Haute Autorité de lutte contre les discriminations et pour l'égalité (Halde), i finalment s'ha integrat l'organisme d'igualtat a la institució del Défenseur des Droits en la seva funció de referent en la lluita contra les discriminacions i promoció de la igualtat (que no té poder per presentar accions col·lectives davant dels tribunals). Tanmateix, en el cas espanyol, i la raó per la qual en aquest treball no combreguem amb aquesta postura, és que la Defensoria del poble només pot conèixer lesions discriminatòries que tinguin una font pública, no privada i, per tant, s'hauria de modificar la Constitució per tal que la defensoria del poble pogués assumir també accions discriminatòries de fonts privades (pàg. 340-341).

L'article 40 de la Llei 15/2022 disposa les funcions de l'Autoritat Independent, entre elles, i que volem destacar per la seva rellevància en la lluita envers la discriminació algorítmica, es troben les següents: c) Iniciar d'ofici o a instància de tercers investigacions sobre l'existència de possibles situacions de discriminació que plantegin especial gravetat o rellevància; d) Exercitar accions judicials en defensa dels drets derivats de la igualtat de tracte i no discriminació; e) Interessar l'actuació de l'administració de l'Estat per sancionar accions o omissions que puguin ser constitutives d'infracció administrativa en aquesta matèria; g) Promoure l'adopció de codis de bones pràctiques en aquesta matèria; h) Col·laborar amb el Defensor del Poble espanyol i les institucions i organismes públics equivalents autonòmics i internacionals; i k) Elaborar, en coordinació amb els òrgans de l'Administració General de l'Estat competents en matèria estadística, informes i estadístiques de caràcter periòdic, promoure estudis sobre igualtat de tracte i no discriminació, així com sobre les formes històriques de discriminació estructural, de les quals han sigut víctimes els grups als quals pretén protegir la llei, dissenyar i mantenir un baròmetre sobre la igualtat de tracte i no discriminació partint d'un sistema d'indicadors i divulgar les activitats, estudis i informes que realitzi¹⁸⁴⁹.

¹⁸⁴⁹ L'article 40 de la Llei 15/2022 descriu les següents funcions: a) garantir la prestació independent de serveis especialitzats d'assistència i orientació a les persones que hagin pogut patir discriminació protegida per la llei, incloent la recepció i tramitació de queixes o reclamacions de les víctimes i activitats de mediació i conciliació com també l'exercici d'accions judicials; b) constituir-se en òrgan de mediació o conciliació, amb decisions vinculants per les parts, sempre amb el seu consentiment i llevat dels casos on hi hagi contingut penal o laboral; c) Iniciar d'ofici o a instància de tercers investigacions sobre l'existència de possibles situacions de discriminació que plantegin especial gravetat o rellevància; d) Exercitar accions judicials en defensa dels drets derivats de la igualtat de tracte i no discriminació; e) Interessar l'actuació de l'administració de l'Estat per sancionar accions o omissions que puguin ser constitutives d'infracció administrativa en aquesta matèria; f) Posar en coneixement del Ministeri Fiscal fets que puguin ser constitutius d'infracció penal; g) Promoure l'adopció de codis de bones pràctiques en aquesta matèria; h) Col·laborar amb el Defensor del Poble espanyol i les institucions i organismes públics equivalents autonòmics i internacionals; i) Emetre dictamen sobre els projectes de disposicions de caràcter general que desenvolupin aquesta llei, així com qualsevol altre que afecti al dret a la igualtat i no discriminació constitucionalment reconegut; j) Informar, amb caràcter preceptiu, sobre l'Estratègia Estatal per a la Igualtat de Tracte i la No Discriminació, així com aquells plans i programes estatals d'especial rellevància en la matèria; k) Elaborar, en coordinació amb els òrgans de l'Administració General de l'Estat competent en matèria estadística, informes i estadístiques de caràcter periòdic, promoure estudis sobre igualtat de tracte i no discriminació, així com sobre les formes històriques de discriminació estructural, de les quals han sigut víctimes els grups als que pretén protegir la llei, dissenyar i mantenir un baròmetre sobre la igualtat de tracte i no discriminació partint d'un sistema d'indicadors i divulgar les activitats, estudis i informes que realitzi; l) Vetllar pel compliment de la normativa reguladora de la igualtat de tracte i no discriminació en l'àmbit de les seves competències així com formular propostes per la seva modificació; m) Informar a instància dels òrgans judicials en els processos jurisdiccionals o del Ministeri Fiscal en les diligències prèvies que versin sobre els drets derivats de la igualtat de tracte i no discriminació; n) Elaborar i proposar al Govern, per la seva aprovació, l'Estatut de l'Autoritat Independent i les seves eventuais modificacions; ñ) Aprovar l'informe anual de les seves activitats, que remetrà al Congrés dels Diputats, al Govern i al Defensor del Poble espanyol; o) Participar en el Foro para la integración social de los inmigrantes; p) Participar en la Comisión Laboral Tripartita de Inmigración, en els termes prevists en la legislació; q)

Iniciar d'ofici o a instància de tercers investigacions sobre l'existència de discriminació, com hem evidenciat anteriorment, pot possibilitar que l'Autoritat independent investigui i sobretot demani i obtingui informació d'empreses, organitzacions o administracions públiques que puguin estar usant sistemes algorítmics potencialment discriminatoris. En aquest sentit, pren rellevància l'article 44.4 de la norma, que disposa que les administracions públiques i els particulars hauran de prestar col·laboració necessària a l'Autoritat Independent quan així ho exigeixi el compliment de la seva funció d'investigació de possibles situacions de discriminació¹⁸⁵⁰ i hauran de proporcionar, quan se'ls requereixi i en termini, tota classe d'informació i dades de què disposin i pugin resultar necessàries pel seu compliment. El termini serà de 10 dies llevat que per la naturalesa d'allò sol·licitat o les circumstàncies del cas es fixi de forma motivada un termini diferent¹⁸⁵¹. El deure de col·laboració i informació inclourà la comunicació de la informació que contingui dades personals de tercers sense el seu consentiment quan resulti estrictament necessari pel compliment de les funcions de l'Autoritat Independent de conformitat amb el que es disposa en l'LOPDGDD i la Ley 12/1989, de 9 de mayo, de la Función Estadística Pública¹⁸⁵². Entre les infraccions greus previstes a la llei s'inclou l'incompliment d'un requeriment administratiu específic, que no constitueixi una exigència formal, formulat per l'òrgan administratiu al qual correspongui l'exercici de les competències necessàries per donar compliment a la llei¹⁸⁵³, de manera que estarà sancionat no respondre a la sol·licitud d'informació de l'Autoritat Independent. Tanmateix, potser s'hauria d'haver inclòs específicament l'omissió del deure de col·laborar amb l'Autoritat Independent, fet que plantejaria de manera més expressa les conseqüències de no donar la informació sol·licitada per part de l'Autoritat Independent¹⁸⁵⁴. Això ens porta a analitzar la funció de l'apartat (e), que suposa interessar

Participar en el Consejo para la Promoción de la Igualdad de Trato y No Discriminación de las Personas por el Origen Racial o Étnico; r) Qualsevol altra que li sigui atribuïda per llei o reglamentàriament.

¹⁸⁵⁰Les prerrogatives que té l'organisme d'igualtat en relació amb la investigació de situacions de discriminació arran de l'article 44.4 de la Llei 15/2022 milloren el que estava establert en l'article 31.1 de la Llei 62/2003, relatiu a l'actuació administrativa contra la discriminació, i que citem seguidament: *“Cuando una autoridad pública, con ocasión del ejercicio de sus competencias, tenga conocimiento de un supuesto de discriminación de los previstos en esta ley, deberá, si es competente, incoar el correspondiente procedimiento administrativo, en el que se podrán acordar las medidas necesarias para investigar las circunstancias del caso y adoptar las medidas oportunas y proporcionadas para su eliminación o, en caso de no serlo, comunicar estos hechos de forma inmediata a la Administración competente, de acuerdo con lo establecido en las leyes administrativas”*.

¹⁸⁵¹ Article 44.3 de la Llei 15/2022.

¹⁸⁵² Article 44.4 de la Llei 15/2022.

¹⁸⁵³ Article 47.3.c de la Llei 15/2022.

¹⁸⁵⁴ REY MARTÍNEZ, F., “O organismo español para a igualdade...”, *op. cit.*, 2021, pàg. 338.

l'actuació de l'Administració de l'Estat per sancionar accions o omissions que puguin ser constitutives d'infracció administrativa en aquesta matèria¹⁸⁵⁵. Aquí s'ha perdut una gran oportunitat per l'Autoritat Independent de guanyar visibilitat ciutadana tal com l'AEPD, atès que la llei disposa que l'Autoritat Independent no pot sancionar per ella mateixa, sinó que ha d'interessar l'actuació de l'Administració de l'Estat per sancionar infraccions administratives. Pot haver diferències de criteri entre l'òrgan sancionador i l'Autoritat Independent i sempre prevaldrà la voluntat de l'òrgan sancionador, debilitant de manera manifesta la posició institucional de l'Autoritat Independent¹⁸⁵⁶. Tanmateix, si més no, pot proposar sancions de manera que si detecta un sistema algorítmic discriminatori podrà interessar que es sancioni l'empresa o administració pública que l'hagi emprat (malgrat aquesta última opció sembla poc probable si com a òrgan independent -en teoria- no pot posar pròpiament les sancions).

En l'apartat (d) de l'article 40 es possibilita a l'Autoritat Independent per exercitar accions judicials en defensa dels drets derivats de la igualtat de tracte i la no discriminació. Com ja hem analitzat extensament en el capítol VII d'aquest treball, aquesta funció creiem que té summa rellevància en la lluita col·lectiva-institucional envers la discriminació algorítmica. Tanmateix, sí que és cert que es podria haver anat més enllà i haver facultat l'Autoritat Independent a defensar la prohibició de discriminació directament davant els tribunals -independentment de l'existència o no d'una víctima identificada-, si més no, en els casos més greus, institucionals i sistèmics. Malgrat que l'Autoritat Independent està legitimada per la defensa de la prohibició de discriminació, sempre ha de ser amb l'autorització de les persones afectades (si no és que es tracta d'interessos difusos), una legitimació de la qual ja disposen molts altres subjectes. En el cas de possibilitar que pogués actuar en qualsevol cas en el qual tingués constància que s'està produint discriminació, s'hauria guanyat una oportunitat que la institució fos de referència en la lluita envers la discriminació¹⁸⁵⁷.

En l'apartat (h) del mateix article es disposa que l'Autoritat independent ha de col·laborar amb el Defensor del Poble espanyol i les institucions i organismes públics equivalents autonòmics i internacionals. És rellevant puntualitzar que en la proposició de llei, l'article 36 preveia la possibilitat que les comunitats autònomes poguessin, per llei, crear

¹⁸⁵⁵ Article 40.e de la Llei 15/2022.

¹⁸⁵⁶ REY MARTÍNEZ, F., "O organismo español para a igualdad...", *op. cit.*, 2021, pàg. 338.

¹⁸⁵⁷ *Ídem*, pàg. 339.

organismes amb la mateixa naturalesa, o atribuir a aquells ja existents, alguna o totes les funcions de l'organisme creat per la llei¹⁸⁵⁸. Aquesta disposició no va mantenir-se en la redacció definitiva, però trobem rellevant apuntar la seva existència, ja que era una bona proposta per tal de donar poders a òrgans territorialment més propers a la ciutadania i incrementar els organismes d'igualtat amb possibilitat de litigar en matèria antidiscriminació. Tot i així, és cert que preservar que existeixi un únic òrgan facilita la seva identificació i visibilitat.

L'adopció de codis de bones pràctiques en aquesta matèria¹⁸⁵⁹ pot, en el cas de la discriminació algorítmica, ajudar les empreses que desenvolupin sistemes algorítmics a tenir en compte la possibilitat que el sistema sigui discriminatori i establir salvaguardes en aquest sentit, sota la guia d'un organisme especialitzat en l'àmbit que té com a prioritat la no discriminació. Finalment, l'elaboració d'informes i estadístiques de caràcter periòdic i la promoció d'estudis sobre igualtat de tracte i no discriminació, així com sobre les formes històriques de discriminació estructural, de les quals han sigut víctimes els grups als quals la llei pretén protegir¹⁸⁶⁰, com ja hem esmentat anteriorment, pot contribuir a crear elements comparatius en el cas que es vulgui desafiar la discriminació indirecta causada per un sistema algorítmic de presa de decisió. Interessa també, en aquest sentit i per evidenciar com la llei sembla apostar per la confecció d'estadístiques com a element clau en la lluita envers la discriminació, l'article 36 de la norma, que estableix que els poders públics elaboraran estudis, memòries o estadístiques, sempre que es refereixin o afectin a aspectes relacionats amb la igualtat de tracte, d'acord amb els estàndards internacionals existents, que permetin un millor coneixement de les causes, extensió, evolució, naturalesa i efectes de la discriminació¹⁸⁶¹.

Cal puntualitzar que a l'Autoritat Independent no se li atorguen en cap cas funcions com a òrgan quasi-jurisdiccional. Només pot prendre decisions pròpies en el context de la mediació o la conciliació (substitutòria al recurs d'alçada i, en el seu cas, de reposició, en relació amb les resolucions de tràmit i actes de tràmit susceptibles d'impugnació) que tindran caràcter vinculant per les parts¹⁸⁶². No s'esmenta en cap cas que pugui prendre

¹⁸⁵⁸ Proposición de Ley integral para la igualdad de trato y la no discriminación. BOCG, núm. 146-1, 29 de enero de 2021. España.

¹⁸⁵⁹ Article 40.g de la Llei 15/2022.

¹⁸⁶⁰ Article 40.k de la Llei 15/2022.

¹⁸⁶¹ Article 36.1 de la Llei 15/2022.

¹⁸⁶² Article 40.b de la Llei 15/2022.

decisions en relació amb les investigacions que dugui a terme -pot, o bé, interessar l'actuació de l'Administració de l'Estat per sancionar infraccions administratives¹⁸⁶³, o bé, posar en coneixement del Ministeri Fiscal els fets constitutius d'infracció penal¹⁸⁶⁴. En canvi, l'AEPD sí que es pot considerar un òrgan quasi-jurisdiccional, atès que els actes i disposicions dictats per la seva Presidència posen fi a la via administrativa, essent recurribles davant la sala Contenciosa-Administrativa de l'Audiència Nacional¹⁸⁶⁵.

Resumidament, l'Autoritat Independent s'adapta als estàndards europeus sobre organismes nacionals d'igualtat, però la seva configuració podria haver estat més àmplia. Quant a la lluita envers la discriminació algorítmica, hauria estat millor que comptés amb capacitat per funcionar com a òrgan quasi-jurisdiccional i que pogués iniciar en el seu nom accions per protegir la igualtat de tracte i discriminació independentment que hi haguessin víctimes identificades o no. Igualment, cal destacar positivament que pugui iniciar d'ofici o a instància de tercers investigacions sobre l'existència de possibles situacions de discriminació que plantegin especial gravetat o rellevància, l'obligació a administracions públiques i particulars de prestar col·laboració necessària, possibilitant l'accés a informació i dades que l'Autoritat Independent necessiti en les seves investigacions, que pugui exercitar accions judicials en defensa dels drets derivats de la igualtat de tracte i no discriminació i que hagi d'elaborar informes i estadístiques de caràcter periòdic.

Després d'haver estudiat la configuració i funcions de l'Autoritat Independent, analitzem com aquest òrgan pot ser complementari i s'haurà de coordinar i cooperar amb altres òrgans de supervisió de sistemes que impliquen algoritmes (a Espanya, l'AEPD i l'AEISA), per assegurar una bona protecció institucional envers la discriminació algorítmica.

4. Complementarietat, coordinació i cooperació entre l'Autoritat Independent, l'AEISA, l'AEPD

¹⁸⁶³ Article 40.e de la Llei 15/2022.

¹⁸⁶⁴ Article 40.f de la Llei 15/2022.

¹⁸⁶⁵ Article 48.6 de l'LOPDGDD.

Allen i Masters plantegen que s'ha de tenir en compte que la tasca de protecció i promoció de la igualtat contra amenaces de les tecnologies de la IA per part dels organismes d'igualtat ha de suposar intrínsecament una col·laboració amb altres organismes i associacions¹⁸⁶⁶, i en el context d'aquest treball abordem principalment la complementarietat, coordinació i cooperació de l'organisme d'igualtat tant amb les agències de protecció de dades, com amb les autoritats de supervisió dels sistemes d'IA. Ho fem ja que són les autoritats que es creen en base a les regulacions estudiades al llarg del treball i tenint en compte que actualment s'està perfilant quin serà el seu rol en la supervisió de la intel·ligència artificial. És evident que l'organisme d'igualtat té l'expertesa en l'àmbit de la discriminació, però no disposa actualment de les eines per detectar-la quan es produeix mitjançant sistemes algorítmics.

Per tant, per acabar l'estudi de la vessant institucional de la protecció envers la discriminació algorítmica analitzem la complementarietat, la coordinació i la cooperació necessàries entre l'AESIA, l'AEPD i l'Autoritat Independent. Per fer-ho, fem un apunt introductori relatiu a les diferències de regulació dels sistemes algorítmics si s'empren en el sector públic i el sector privat i si això afecta l'actuació de les autoritats de supervisió envers la discriminació algorítmica (4.1), analitzem seguidament les funcions que s'atorguen a l'AESIA (4.2) i les funcions que s'atorguen a l'AEPD (4.3). Finalment estudiem com la complementarietat, la cooperació i la coordinació entre les tres autoritats pot beneficiar la lluita envers la discriminació algorítmica (4.4).

4.1. Diferències de regulació dels usos dels sistemes algorítmics en el sector públic i privat a efectes de l'actuació de les autoritats de supervisió envers la discriminació algorítmica

Les normes que s'apliquen quan els sistemes algorítmics són emprats pels poders públics i quan són emprats per empreses privades es regulen de diferent manera i aquesta diferència de regulació pot afectar l'actuació de l'autoritat de control i supervisió dels sistemes d'IA quan ens referim a la discriminació algorítmica. Tanmateix, com s'analitza

¹⁸⁶⁶ ALLEN, R. i MASTERS, D., *Regulating for an equal AI...*, op. cit., 2020, pàg. 11 i 12.

en aquest apartat, actualment la regulació de l'ús de sistemes algorítmics per part de poders públics no és molt més exigent que aquella que s'estableix en l'actuació d'actors privats. A més, malgrat que les actuacions del sector públic es regeixen per principis més estrictes que les actuacions del sector privat (p. e. principis de bona administració, interdicció d'arbitrarietat i publicitat de les normes), la prohibició de discriminació opera de la mateixa manera per als dos sectors.

En l'àmbit del sector públic el control de la igualtat queda molt lligat i se superposa amb el compliment d'altres mandats constitucionals: l'activitat administrativa no pot ser arbitrària, ha d'estar motivada amb suficiència i racionalitat jurídica (art. 9.3 i 24 de la Constitució Espanyola), ha de sotmetre's a la legalitat i al mandat d'objectivitat (art. 103 de la Constitució Espanyola) i el control inclourà que no es doni un desviament de poder (art. 106 de la Constitució Espanyola)¹⁸⁶⁷.

L'ordenament jurídic espanyol no considera els algoritmes emprats en l'actuació dels poders públics com a normes -Boix Palop ha defensat que haurien de considerar-se normes reglamentàries¹⁸⁶⁸-, sinó que es consideren, segons l'article 41 de la Ley 40/2015, de 1 de octubre, de Régimen Jurídico del Sector Público (d'ara en endavant, Llei 40/2015), actuacions administratives. En el cas de considerar-se normes, l'article 9.3 de la Constitució Espanyola en garantiria la seva publicitat¹⁸⁶⁹.

L'article 41 de la Llei 40/2015 estableix en el seu primer apartat que s'entén per actuació administrativa automatitzada qualsevol acte o actuació realitzada íntegrament a través de mitjans electrònics per una Administració Pública en el marc d'un procediment administratiu i en el qual no hagi intervingut de forma directa un/a treballador/a públic/a¹⁸⁷⁰. Segons aquesta definició, en l'ordenament jurídic espanyol, els algoritmes o programes que prenen decisions administratives es consideren actuacions molt singularitzades que requereixen una explicabilitat suficient com per passar el filtre en termes de motivació segons s'estableix tant l'article 9.3 de la Constitució Espanyola, amb la prohibició d'arbitrarietat dels poders públics, com l'article 35 de la Llei 39/2015, amb

¹⁸⁶⁷ COTINO HUESO, L., "Discriminación, sesgos e igualdad de la inteligencia artificial...", *op. cit.*, 2023, pàg. 285 i 286.

¹⁸⁶⁸ Vegeu: BOIX PALOP, A., "Los algoritmos son reglamentos: la necesidad de extender...", *op. cit.*, 2020, pàg. 223-270.

¹⁸⁶⁹ Article 9.3 de la Constitució Espanyola.

¹⁸⁷⁰ Article 41.1 de la Llei 40/2015.

l'obligació de motivació de certs actes administratius¹⁸⁷¹.

S'ha criticat la redacció de l'article 41 de la Llei 40/2015 ja que limita la consideració d'actuació administrativa automatitzada i no s'aplicarà en els casos en què la intervenció sigui auxiliar, complementària o prèvia a la decisió final, per molt que *de facto* sigui determinant¹⁸⁷². Cerrillo i Martínez apunta que l'ús d'algoritmes en la presa de decisions en l'àmbit públic canvia la manera com aquestes es generen i, en conseqüència, com es motiven els actes administratius, implicant això -malgrat que actualment no existeixi una norma explícita en aquest sentit- que s'hagi d'explicar com s'ha arribat a la decisió, quin ha estat el procés, quines dades s'han tingut en compte i quins objectius es persegueixen¹⁸⁷³. En el segon apartat de l'article 41 de la Llei 40/2015, s'estableix que en cas d'una actuació administrativa automatitzada s'haurà d'establir prèviament l'òrgan o òrgans competents, segons els casos, per la definició d'especificitats, programació, manteniment, supervisió i control de qualitat i, en el seu cas auditoria del sistema de informació i el seu codi font, havent-se d'establir també l'òrgan responsable a efectes d'impugnació de l'actuació¹⁸⁷⁴. Les obligacions que imposa aquest article han estat criticades atesa la seva insuficiència.

L'article 23 de la Llei 15/2022, centrant-se en els efectes discriminatoris de l'ús de sistemes algorítmics en el sector públic, estableix que les administracions públiques afavoriran la implementació de mecanismes per tal que els algoritmes involucrats en la presa de decisions que s'utilitzin en l'administració pública tinguin en compte criteris de minimització de biaixos, transparència i rendició de comptes, com per exemple, avaluacions d'impacte que determinin el possible biaix discriminatori¹⁸⁷⁵. També es disposa que les administracions públiques prioritzaran la transparència en el disseny dels algoritmes i la implementació i capacitat d'interpretació de les decisions adoptades pels mateixos¹⁸⁷⁶ i promouran un segell de qualitat¹⁸⁷⁷. Aquest article deixa marge d'apreciació a les administracions públiques per implementar els mecanismes enumerats en l'article a

¹⁸⁷¹ BOIX PALOP, A. i SORIANO ARNANZ, A. "Transparencia y control del uso de la inteligencia artificial por las administraciones públicas", a BALAGUER CALLEJÓN, F. i COTINO HUESO, L. (coords), *Derecho Público de la Inteligencia Artificial*, Zaragoza, Fundación Manuel Giménez Abad, 2023, pàg. 267-268.

¹⁸⁷² *Ídem, ibídem.*

¹⁸⁷³ CERRILLO I MARTÍNEZ, A., "Com obrir les caixes negres de les administracions públiques? Transparència i rendició de comptes en l'ús dels algoritmes", *Revista Catalana de Dret Públic*, núm. 58, 2019, pàg. 21 i 22.

¹⁸⁷⁴ Article 41.2 de la Llei 40/2015.

¹⁸⁷⁵ Article 23.1 de la Llei 15/2022.

¹⁸⁷⁶ Article 23.2 de la Llei 15/2022.

¹⁸⁷⁷ Article 23.4 de la Llei 15/2022.

través de l'ús dels verbs com “afavorir” o “prioritzar”¹⁸⁷⁸. Tanmateix, l'Autoritat Independent podria supervisar si efectivament les administracions públiques quan usen sistemes algorítmics han seguit els criteris establerts en aquesta norma i, si no fos així, requerir-les per tal que ho fessin.

També cal destacar l'apartat 6 de l'article XVIII¹⁸⁷⁹ i l'apartat 2 de l'article XXV¹⁸⁸⁰ de la Carta de Drets Digitals, que com ja hem esmentat en l'apartat 3.1 del capítol V d'aquest treball, plantegen certes condicions de transparència en relació amb els instruments de la intel·ligència artificial en el marc de l'actuació administrativa. Així i tot, la carta no té caràcter normatiu.

En la PRIA es preveuen tant funcions públiques en les quals està prohibit l'ús de sistemes d'IA (p. e. avaluar o classificar per part de les autoritats públiques la fiabilitat de les persones físiques de forma que en resulti una classificació social¹⁸⁸¹), com funcions públiques en les quals l'ús de sistemes de IA es considera de risc alt (p. e. sistemes d'IA emprats per assumptes relacionats amb l'aplicació de la llei¹⁸⁸², la gestió de la migració, l'asil i el control fronterer¹⁸⁸³ o l'administració de justícia i els processos democràtics¹⁸⁸⁴).

En relació amb la PRIA, també cal esmentar que l'article 64.3 i el considerant 79 de la norma estipulen que les autoritats o organismes públics nacionals que supervisin l'aplicació del dret de la UE que protegeix els drets fonamentals, inclosos els organismes d'igualtat, també tindran accés a la documentació tècnica¹⁸⁸⁵ que es creï en virtut de la PRIA¹⁸⁸⁶. Cotino Hueso planteja que en relació amb els usos públics de sistemes d'IA, aquesta documentació tècnica en possessió de les autoritats competents es tractarà d'informació pública¹⁸⁸⁷ a la qual es podrà accedir a través del dret d'accés a informació

¹⁸⁷⁸ COTINO HUESO, L., “Discriminación, sesgos e igualdad de la inteligencia artificial...”, *op. cit.*, 2023, pàg. 328. L'autor estableix que la llei utilitza termes molt suaus però, tot i així, es possible que generin un important efecte preventiu de la discriminació algorítmica a la pràctica.

¹⁸⁷⁹ Article XVIII.6 de la Carta de Drets Digitals Espanyola.

¹⁸⁸⁰ Article XXV.2 de la Carta de Drets Digitals Espanyola.

¹⁸⁸¹ Article 5.1.c de la PRIA.

¹⁸⁸² Apartat 6 de l'Annex III de la PRIA.

¹⁸⁸³ Apartat 7 de l'Annex III de la PRIA.

¹⁸⁸⁴ Apartat 8 de l'Annex III de la PRIA.

¹⁸⁸⁵ Article 11 de la PRIA.

¹⁸⁸⁶ Article 64.3 i considerant 79 de la PRIA.

¹⁸⁸⁷ L'article 13 de la Llei 19/2013 disposa que: *Se entiende por información pública los contenidos o documentos, cualquiera que sea su formato o soporte, que obren en poder de alguno de los sujetos incluidos en el ámbito de aplicación de este título y que hayan sido elaborados o adquiridos en el ejercicio de sus funciones.*

pública¹⁸⁸⁸.

La manca d'una regulació més exhaustiva s'ha intentat pal·liar a través de l'aplicació de mesures que es preveuen en l'RGPD¹⁸⁸⁹ i la LOPDGDD, intentant que els poders públics, si més no, se sotmetin als mateixos controls que s'imposen en l'àmbit privat¹⁸⁹⁰. Tanmateix, una excepció a la prohibició de l'article 22 de l'RGPD del dret a no ser objecte a decisions individuals automatitzades, enumerada en el quart apartat de l'article, és que el tractament sigui necessari per raons d'un interès públic essencial¹⁸⁹¹, que és el que fa l'article 41.1 de la Llei 40/2015¹⁸⁹², establint l'article 41.2 de la mateixa norma mesures adequades i específiques per a protegir els interessos i drets fonamentals de la persona interessada¹⁸⁹³.

També s'han de referenciar iniciatives autonòmiques per limitar i controlar l'ús de sistemes algorítmics en l'administració pública com el Decreto-ley 2/2023, de 8 de marzo, de medidas urgentes de impulso a la inteligencia artificial en Extremadura¹⁸⁹⁴, en el qual, en l'article 11 s'estableix que quan s'adoptin actes administratius mitjançant sistemes d'intel·ligència artificial en el marc d'un procediment administratiu, a més dels requisits prevists en l'article 41 de la Llei 40/2015, es donarà la deguda publicitat del mecanisme de decisió, de les prioritats assignades en el procediment d'avaluació i de la presa de decisions, així com totes les dades que puguin impactar en el seu contingut. També, en l'article 12, en el primer apartat, s'estableix que les normes que regulin els procediments administratius faran referència expressa a l'impacte de l'ús d'aquests sistemes en la prestació de serveis públics que, en el seu cas, suportin la presentació de sol·licituds, declaracions responsables o comunicacions, la comprovació o verificació dels requisits de les persones interessades, així com la presa de decisions. En el segon apartat s'estableix que es deixarà constància en l'Inventari d'Informació Administrativa, sobre les

¹⁸⁸⁸ COTINO HUESO, L. "Qué concreta transparencia e información de algoritmos...", *op. cit.*, 2023, pàg. 57.

¹⁸⁸⁹ VESTRI, G., "La inteligencia artificial ante el desafío de la transparencia algorítmica. una aproximación desde la perspectiva jurídico-administrativa", *Revista Aragonesa de Administración Pública*, núm. 56, 2021, pàg. 387-388; CERRILLO I MARTÍNEZ, A., "¿Son fiables las decisiones de las Administraciones públicas adoptadas por algoritmos?", *European Review of Digital Administration & Law - Erdal*, vol. 1, núm. 1-2, 2020, pàg. 17-36.

¹⁸⁹⁰ BOIX PALOP, A. i SORIANO ARNANZ, A. "Transparencia y control del uso...", *op. cit.*, 2023, pàg. 271 i 272.

¹⁸⁹¹ Article 22.4 de l'RGPD en relació amb l'article 9.2.g de la mateixa norma.

¹⁸⁹² Article 41.1 de la Llei 40/2015.

¹⁸⁹³ Article 41.2 de la Llei 40/2015.

¹⁸⁹⁴ Decreto-ley 2/2023, de 8 de marzo, de medidas urgentes de impulso a la inteligencia artificial en Extremadura. DOE, núm. 48, de 10 de marzo de 2023.

validacions realitzades per l'òrgan responsable dels procediments administratius respecte el procés lògic dissenyat per la realització dels actes, els riscos que implica i qualsevol altre aspecte que garanteixi els drets de les persones interessades. De la mateixa manera, la Llei 1/2022, de 13 d'abril, de la Generalitat, de Transparència i Bon Govern de la Comunitat Valenciana¹⁸⁹⁵, en el seu article 16.1.1 estableix que les administracions públiques hauran de publicar “*la relació de sistemes algorítmics o d'intel·ligència artificial que tinguen impacte en els procediments administratius o la prestació dels serveis públics amb la descripció de manera comprensible del seu disseny i funcionament, el nivell de risc que impliquen i el punt de contacte al qual poder adreçar-se en cada cas, d'acord amb els principis de transparència i explicabilitat*”. Aquesta norma, a diferència de l'article 41.2 de la Llei 40/2015, obliga a incloure en el registre tots els algoritmes emprats que directa o indirectament afectin o condicionin la decisió administrativa final, sense limitar-se només a les decisions exclusivament algorítmiques¹⁸⁹⁶.

Com exposen Boix Palop i Soriano Aranz, malgrat que l'actuació de l'Administració pública ha de ser sempre estrictament respectuosa amb els drets de la ciutadania i quedar subjecte a més obligacions i controls que les relacions entre particulars atesos els poders i potestats que exerceix el poder públic sobre la ciutadania¹⁸⁹⁷, la reacció del poder legislatiu quan els algoritmes s'han emprat per part d'empreses privades ha estat molt més emfàtica, imposant obligacions i atorgant drets a la ciutadania enfront els usos d'algoritmes o sistemes d'IA que els afecten que són majors que les obligacions que l'ordenament jurídic espanyol i europeu preveu pels casos en què poders públics entren aquesta tecnologia¹⁸⁹⁸. Per evidenciar-ho, posen l'exemple del Reglamento de Mercados Digitales, que en el seu article 21.3 possibilita expressament a la Comissió Europea, per mitjà d'una sol·licitud o decisió, exigir l'accés a qualsevol dada o algoritme de les empreses i informació sobre les proves, així com sol·licitar explicacions sobre aquesta informació¹⁸⁹⁹, o la Ley 12/2021, de 28 de septiembre, por la que se modifica el texto refundido de la Ley del Estatuto de los Trabajadores, aprobado por el Real Decreto Legislativo 2/2015, de 23 de octubre, para garantizar los derechos laborales de las

¹⁸⁹⁵ Llei 1/2022, de 13 d'abril, de la Generalitat, de Transparència i Bon Govern de la Comunitat Valenciana. DOGV, núm. 9323, de 22 d'abril de 2022.

¹⁸⁹⁶ BOIX PALOP, A. i SORIANO ARANZ, A. “Transparencia y control del uso...”, *op. cit.*, 2023, pàg. 276.

¹⁸⁹⁷ *Ídem*, pàg. 249.

¹⁸⁹⁸ *Ídem*, pàg. 253.

¹⁸⁹⁹ Article 21.3 del Reglamento de Mercados Digitales.

personas dedicadas al reparto en el ámbito de plataformas digitales¹⁹⁰⁰, que introduceix l'apartat d) en l'article 64.4 de la Ley del Estatuto de los Trabajadores que disposa que el Comitè d'Empresa té dret a ser informat sobre els paràmetres, regles i instruccions en què es basen els algorismes i sistemes d'intel·ligència artificial que afecten la presa de decisions)¹⁹⁰¹.

Específicament en relació amb la transparència, en l'administració pública s'estan emprant sistemes algorítmics, com per exemple VioGen, Riscanvi (a Catalunya), o BOSCO, quan no es coneix exactament el seu funcionament ni la ponderació exacta que realitzen per tractar-se d'un sistema algorítmic de suport¹⁹⁰². Justament, com ja hem referenciat extensament en el treball, en relació amb l'algoritme BOSCO, Civio ha plantejat judicialment la seva impugnació atesa la negativa del Consell de Transparència i Bon Govern de proporcionar l'accés al codi font de l'algoritme¹⁹⁰³. El respecte del principi de transparència per part de les administracions públiques és un element essencial del control eficaç dels poders públics per part de la ciutadania, i tanmateix, a la pràctica aquest principi es troba amb moltes limitacions quan ens referim a l'ús de sistemes algorítmics¹⁹⁰⁴, sobretot perquè actualment només es considera actuació administrativa automatitzada aquella que és únicament automatitzada (i igualment la norma que la regula -l'article 41 de la Llei 40/2015- no estableix requisits de transparència), i perquè malgrat estar l'actuació dels poders públics sota un escrutini més exigent de transparència, encara no hi ha una regulació europea o nacional imperativa que obligui a revelar informació rellevant del funcionament del sistema algorítmic (vegeu, per exemple, la disputa en el cas BOSCO). Cotino Hueso afirma que és necessari projectar garanties de transparència

¹⁹⁰⁰ Ley 12/2021, de 28 de septiembre, por la que se modifica el texto refundido de la Ley del Estatuto de los Trabajadores, aprobado por el Real Decreto Legislativo 2/2015, de 23 de octubre, para garantizar los derechos laborales de las personas dedicadas al reparto en el ámbito de plataformas digitales. BOE, núm. 233, de 29 de septiembre de 2021. España.

¹⁹⁰¹ BOIX PALOP, A. i SORIANO ARNANZ, A. "Transparencia y control del uso...", *op. cit.*, 2023, pàg. 252 i 253.

¹⁹⁰² *Ídem*, pàg. 255.

¹⁹⁰³ Vegeu: CONSEJO DE TRANSPARENCIA Y BUEN GOBIERNO, Resolución 701/2018, de 18 de febrero de 2019. Disponible a: <https://civio.app.box.com/s/0r9a58a10pdp2t08lnemlmc22sd7t31>. Darrera visita el 2 de setembre de 2023.

¹⁹⁰⁴ JIMÉNEZ-CASTELLANOS BALLESTEROS, I., "Decisiones automatizadas y transparencia administrativa: nuevos retos para los derechos fundamentales", *Revista Española de la Transparencia (RET)*, núm. 16, 2023, pàg. 212 i 213.

i participació a través d'una nova normativa i actualització de la interpretació d'aquella existent¹⁹⁰⁵.

En definitiva, l'actuació de les autoritats que supervisin l'ús de sistemes algorítmics en la presa de decisions en contextos socials (p. e. AESIA, AEPD o l'Autoritat Independent), que són independents i poden actuar en relació a sistemes algorítmics emprats tant en el sector públic com privat, actualment no difereix molt si el sistema algorítmic s'ha emprat en el sector públic o en el sector privat.

4.2. L'AESIA i la protecció envers la discriminació algorítmica

Des de la doctrina s'ha especulat sobre les atribucions que hauria de tenir una autoritat que supervisés l'ús d'algoritmes. Per exemple, Roig ha plantejat que l'autoritat tingui com a funció principal moderar i gestionar plataformes digitals de participació¹⁹⁰⁶, o Simón Catellano ha proposat que l'autoritat tingui un enfocament específic en els principis que pretenen assolir una intel·ligència artificial ètica i confiable, que exerceixi funcions de certificació, inspecció, sanció i elaboració de recomanacions, avisos i guies sobre la matèria¹⁹⁰⁷. També, Mantelero exposa que l'adopció d'una avaluació de la tecnologia centralitzada duta a terme per autoritats de supervisió nacionals *ad hoc* pot proveir de guies útils per al desenvolupament tecnològic i fixar línies vermelles, complementant-se necessàriament amb una avaluació específica cas per cas de l'impacte de cada aplicació desenvolupada¹⁹⁰⁸.

En el context estatunidenc, Tutt va plantejar que l'esquema regulador de la U.S. Food and Drug Administration (d'ara en endavant, FDA) podia ser adequat per regular els sistemes algorítmics, justificant-ho en què les preocupacions que impulsen una regulació dels medicaments són semblants a les preocupacions que haurien d'impulsar una

¹⁹⁰⁵ COTINO HUESO, L., “Los usos de la inteligencia artificial en el sector público...”, *op. cit.*, 2023, pàg. 228 i 229.

¹⁹⁰⁶ ROIG, A., *Las Garantías frente a las decisiones automatizadas...*, *op. cit.*, 2020, pàg. 237 i 238;

¹⁹⁰⁷ SIMÓN CASTELLANO, P., “Allende una teoría general de las garantías jurídicas para una inteligencia artificial confiable”, a BALAGUER CALLEJÓN, F. I COTINO HUESO, L. (coords), *Derecho Público de la Inteligencia Artificial*, Zaragoza, Fundación Manuel Giménez Abad, 2023, pàg. 138.

¹⁹⁰⁸ MANTELERO, A., *Beyond Data. Human Rights, Ethical and Social Impact Assessment in AI*, *op. cit.*, 2022, pàg. 48.

regulació dels algoritmes¹⁹⁰⁹. D'aquesta manera, l'autoritat hauria de comptar amb capacitat per establir estàndards (p. e. sobre la classificació dels algoritmes segons la seva complexitat; sobre criteris de disseny, prova i rendiment, o sobre la distribució de la responsabilitat en cas de dany causat per l'algoritme), per actuar com a supervisora del compliment de la regulació de dret tou (p. e. si s'assoleixen els requeriments de transparència), i per imposar restriccions al desenvolupament de cert tipus d'algoritmes d'aprenentatge automàtic o per requerir aprovació prèvia a la comercialització d'un sistema abans que es pugui implementar¹⁹¹⁰.

Actualment, a escala europea, és en la PRIA on trobem la primera obligació als estats membres de designar una autoritat competent per supervisar sistemes d'intel·ligència artificial. L'article 59 de la PRIA disposa que cada estat membre designarà autoritats nacionals competents amb la finalitat de garantir l'aplicació i execució del Reglament de la IA¹⁹¹¹, i una autoritat nacional de supervisió que actuarà com a autoritat notificant i com a autoritat de vigilància de mercat¹⁹¹². Aquestes autoritats disposaran permanentment de suficient personal, les competències i coneixements tècnics dels quals inclouran un coneixement profund de les tecnologies de la intel·ligència artificial, dades i computació de dades, riscos pels drets fonamentals, la salut i la seguretat i coneixements sobre normes i requisits legals vigents¹⁹¹³. El punt de partida que planteja la PRIA quant a l'autoritat de supervisió que exigeix crear als estats membres, no dista molt d'aquella plantejada per Tutt, amb funcions de supervisió i inclús d'aprovació prèvia per certs sistemes determinats abans de la seva introducció en el mercat i creació d'estàndards i criteris d'actuació (com per exemple fan actualment les autoritats de protecció de dades¹⁹¹⁴). A més, el Parlament Europeu en les seves propostes d'esmena al text de la Comissió Europea inclou que les persones perjudicades per un sistema d'IA tinguin dret a presentar una queixa davant l'AESIA¹⁹¹⁵, de manera que si aquesta esmena es recull en el text final,

¹⁹⁰⁹ TUTT, A., "An FDA for algorithms", *Administrative Law Review*, vol. 69, núm. 1, 2017, pàg. 119-120.

¹⁹¹⁰ *Ídem*, pàg. 107-110.

¹⁹¹¹ Article 59.1 de la PRIA.

¹⁹¹² Article 59.2 de la PRIA.

¹⁹¹³ Article 59.4 de la PRIA.

¹⁹¹⁴ Vegeu, per exemple: AEPD, *Orientaciones para la realización de una evaluación de impacto para la protección de datos en el desarrollo normativo*, Madrid, 18 de abril de 2023; AUTORITAT CATALANA DE PROTECCIÓ DE DADES, *Dictamen en relació amb la consulta formulada per un organisme en relació amb la instal·lació de càmeres de videovigilància pel control del trànsit en el casc urbà*, Barcelona, 11 de febrer de 2021.

¹⁹¹⁵ EUROPEAN PARLIAMENT, *Amendments adopted by the European Parliament on 14 June 2023...*, *op. cit.*, 14 June 2023.

l'AESIA també tindrà la funció de gestionar les queixes de les persones perjudicades.

A Espanya, la disposició addicional centèsima trigèsima de la Ley 22/2021, de 28 de diciembre, de Presupuestos Generales del Estado para el año 2022¹⁹¹⁶ creava l'AESIA, configurada com a Agència Estatal dotada de personalitat jurídica pública, patrimoni propi i autonomia de gestió amb potestat administrativa¹⁹¹⁷. Es garanteix en el segon apartat d'aquesta disposició addicional que l'AESIA actuarà amb plena independència orgànica i funcional de les Administracions Públiques¹⁹¹⁸, de forma objectiva, transparent i imparcial, prenent mesures destinades a la minimització de riscos significatius sobre la seguretat i la salut de les persones, així com sobre els seus drets fonamentals, que puguin derivar-se de l'ús de la intel·ligència artificial¹⁹¹⁹. Segons el Real Decreto 729/2023, de 22 de agosto, por el que se aprueba el Estatuto de la Agencia Española de Supervisión de Inteligencia Artificial (d'ara en endavant, Estatut de l'AESIA), un dels seus principis d'actuació és la independència tècnica basa de la capacitació, especialització, professionalitat i responsabilitat individual del personal al servei de l'Agència (...), actuant amb plena autonomia¹⁹²⁰. Tanmateix està adscrita al Ministeri d'Assumptes Exteriors i Transformació Digital i no s'especifica quin serà el procés per anomenar la persona que ostentarà la seva presidència¹⁹²¹. L'Estatut de l'AESIA, estableix les competències que exercirà¹⁹²², i integra tant competències de control *ex ante* com de

¹⁹¹⁶ Ley 22/2021, de 28 de diciembre, de Presupuestos Generales del Estado para el año 2022. BOE, núm. 312, de 29/12/2021. España.

¹⁹¹⁷ Apartat 1 de la Disposició addicional centèsima trigèsima de la Ley 22/2021, de 28 de diciembre, de Presupuestos Generales del Estado para el año 2022.

¹⁹¹⁸ Tot i així, en l'apartat quart de la mateixa disposició s'estableix que es trobarà adscrita a la Secretaria d'Estat de Digitalització i Intel·ligència Artificial, dins del Ministeri d'Assumptes Econòmics i Transformació Digital.

¹⁹¹⁹ Apartat 2 de la Disposició addicional centèsima trigèsima de la Ley 22/2021, de 28 de diciembre, de Presupuestos Generales del Estado para el año 2022.

¹⁹²⁰ Article 8.b del Real Decreto 729/2023, de 22 de agosto, por el que se aprueba el Estatuto de la Agencia Española de Supervisión de Inteligencia Artificial.

¹⁹²¹ Article 1.1 del Real Decreto 729/2023, de 22 de agosto, por el que se aprueba el Estatuto de la Agencia Española de Supervisión de Inteligencia Artificial.

¹⁹²² Article 10 del Real Decreto 729/2023, de 22 de agosto, por el que se aprueba el Estatuto de la Agencia Española de Supervisión de Inteligencia Artificial. Per exemple: a) la promoció d'entorns de prova (...); b) el suport al desenvolupament i ús de sistemes d'intel·ligència artificial des d'una doble perspectiva ambiental i sostenible (...); d) l'ajuda a l'enfortiment de la confiança en la tecnologia i aplicació de la intel·ligència artificial a través d'un marc de certificació voluntària per a entitats privades (...); e) la instrumentalització dels mecanismes d'identificació de tendències i avaluació de l'impacte social en matèria d'intel·ligència artificial (...); f) l'alineament i coordinació amb iniciatives de tercers relacionades amb l'aplicació dels sistemes d'intel·ligència artificial (...); g) la creació de coneixement, formació i difusió en relació amb la intel·ligència artificial ètica i humanista (...); h) la dinamització del mercat per potenciar iniciatives i practiques innovadores i transformadores (...); i) l'impuls de la col·laboració públicoprivada (...); j) l'ajuda en l'execució de programes (...); k) la supervisió dels sistemes d'intel·ligència artificial per garantir el compliment de la normativa (...) que es realitzarà amb perspectiva de gènere; l) Qualsevol altra funció que pogués ser atribuïda a l'Agència (...).

control *ex post*. Aquesta dualitat s'explica perfectament amb la conclusió a la qual va arribar l'FDA en el seu pla d'acció per actualitzar els seus protocols per a la intel·ligència artificial, aprenentatge automàtic i software com a dispositius mèdics, de gener de 2021, on va determinar que el seu paradigma tradicional de regulació de medicaments no estava dissenyat per a la intel·ligència artificial i les tecnologies d'aprenentatge automàtic atesa la seva capacitat d'adaptació¹⁹²³. És a dir, el seu paradigma tradicional que requereix als participants en la indústria d'assolir certs nivells de seguretat i eficàcia abans de rebre la seva aprovació i permetre que els medicaments arribin al públic general¹⁹²⁴, no es pot aplicar a dispositius mèdics que integren IA, ja que malgrat rebin l'aprovació abans de la seva introducció al mercat, la IA pot fer que el dispositiu que va assolir certs nivells de seguretat i eficàcia en el moment de ser aprovat, posteriorment ja no els assoleixi per la seva capacitat d'evolució i canvi.

Quant a les competències, és d'especial rellevància la que es descriu en l'article 10.1.c de l'Estatut de l'AESIA, de suport al desenvolupament i l'ús de sistemes d'intel·ligència artificial amb perspectiva de gènere, incorporant el principi d'igualtat entre dones i homes en el seu disseny i execució i promovent la realització d'avaluacions d'impacte capaces d'identificar possibles biaixos discriminatoris per qualsevol dels motius prohibits per l'ordenament jurídic, per tal d'eliminar els biaixos discriminatoris de qualsevol tipus i, en particular, aquells de gènere i ètnics-racials¹⁹²⁵. També cal destacar que en relació amb la cooperació amb altres agències i autoritats, l'Estatut estableix una línia d'actuació clara. L'article 4.3.c estableix que una de les finalitats de l'AESIA serà la col·laboració i coordinació amb altres autoritats, nacionals i supranacionals de supervisió de la intel·ligència artificial¹⁹²⁶. També l'article 8.f estableix que un dels principis d'actuació de l'Agència serà la cooperació interinstitucional, entesa com la recerca de sinergies en la col·laboració amb altres administracions públiques, agents o institucions, públiques o

¹⁹²³ U.S. FOOD AND DRUG ADMINISTRATION (FDA), "Artificial intelligence machine learning and (AI/ML) software as a medical Device (SaMD) action plan", 2021. Disponible a: <https://www.fda.gov/medical-devices/software-medical-device-samd/artificial-intelligence-and-machine-learning-software-medical-device> Darrera visita el 2 de setembre de 2023.

¹⁹²⁴ SOLOW-NIEDERMAN, A., "Administering Artificial Intelligence", *op. cit.*, 2020, pàg. 651-652; MERRILL, R. A., "The Architecture of Government Regulation of Medical Products", *Virginia Law Review*, vol. 82, núm. 8, 1996, pàg. 1764-1767. L'autor precisa, i és interessant en relació amb la regulació establerta en la PRIA per la majoria de sistemes d'IA, que l'FDA va passar d'un sistema previ al mercat de notificació a un sistema previ al mercat d'aprovació.

¹⁹²⁵ Article 10.1.c del Real Decreto 729/2023, de 22 de agosto, por el que se aprueba el Estatuto de la Agencia Española de Supervisión de Inteligencia Artificial.

¹⁹²⁶ Article 4.3.c del Real Decreto 729/2023, de 22 de agosto, por el que se aprueba el Estatuto de la Agencia Española de Supervisión de Inteligencia Artificial.

privades, nacionals o internacionals, per al foment del coneixement en tots els seus àmbits¹⁹²⁷.

Com ja hem vist en l'apartat 4.2 del capítol VI, la majoria de sistemes de risc alt podran fer una avaluació de conformitat basada en el control intern¹⁹²⁸, requerint només la implicació d'organismes notificats en l'àrea de la identificació biomètrica i la categorització de persones naturals (punt 1 de l'annex III de la PRIA)¹⁹²⁹. Això ens fa pensar que aquesta autoritat, en tant que òrgan de control *ex ante* a la posada en funcionament dels sistemes, a part de vetllar per tal que no s'usin sistemes prohibits (que poden causar també discriminació algorítmica), en el cas dels sistemes de risc alt, només podrà abastar els casos més evidents de discriminació algorítmica, atesa la seva implicació diluïda en el procés d'introducció en el mercat. Els casos de discriminació algorítmica evident es podran detectar a través de l'examen i avaluació de la documentació tècnica que hagin presentat els proveïdors del sistema de risc alt¹⁹³⁰ (que ha de contenir la precisió del sistema algorítmic segons diferents grups de població¹⁹³¹), malgrat que s'ha de tenir en compte que si el sistema no és de risc alt no caldrà presentarla. L'exigència d'aquesta documentació tècnica pot llegir-se com un intent de contenir els problemes d'interpretabilitat dels algoritmes¹⁹³². Per tant, aquesta autoritat podrà detectar l'existència de discriminació algorítmica arran de la documentació tècnica, arran del desenvolupament dels sistemes d'IA en els espais de prova que proposa la PRIA¹⁹³³ o arran de la supervisió de l'ús del tractament de categories especials de dades per mitigar els biaixos algorítmics¹⁹³⁴.

Cal puntualitzar que en la PRIA els sistemes d'IA són classificats segons el risc, i aquest ve completament determinat d'avançada sense tenir en compte el context en el qual s'implementarà i, arran de la interacció amb el context, el nivell de risc canviarà. Això ha estat estudiat en el cas de la discriminació algorítmica, on inicialment pot semblar que el sistema funciona correctament però en aplicar-lo en un context determinat, causa

¹⁹²⁷ Article 8.f del Real Decreto 729/2023, de 22 de agosto, por el que se aprueba el Estatuto de la Agencia Española de Supervisión de Inteligencia Artificial.

¹⁹²⁸ Article 43.2 de la PRIA.

¹⁹²⁹ VEALE, M. i ZUIDERVEEN, F., "Demystifying the Draft EU...", *op. cit.*, 2021, pàg. 106.

¹⁹³⁰ Article 11 de la PRIA.

¹⁹³¹ Annex IV de la PRIA.

¹⁹³² Vegeu, en relació amb els problemes d'interpretabilitat: SOLOW-NIEDERMAN, A., "Administering Artificial Intelligence", *op. cit.*, 2020, pàg. 657-659.

¹⁹³³ Article 53 de la PRIA.

¹⁹³⁴ Article 10.5 de la PRIA.

discriminació¹⁹³⁵. En aquest sentit, interessa sobretot per part de l'AESIA, en tant que autoritat supervisora de la intel·ligència artificial, la introducció d'una perspectiva de control cíclic i periòdic dels sistemes d'IA.

En el cas de detectar discriminació algorítmica posteriorment a l'entrada d'un sistema algorítmic en el mercat intern, l'AESIA podrà sancionar a qui ha proveït el sistema d'IA per incompliment dels requisits establerts en el Reglament de la IA¹⁹³⁶. Seria interessant que s'establís tant per a administracions públiques com per a particulars un deure de col·laboració amb l'AESIA per proporcionar-li tota la informació que sigui necessària per a dur a terme les seves tasques d'investigació, existint la possibilitat de sanció en cas d'incompliment, de la mateixa manera que tenen aquesta potestat l'Autoritat Independent -article 44.4 de la Llei 15/2022- i l'AEPD -article 52 de l'LOPDGDD-.

La PRIA no atorga drets a les persones perjudicades per un sistema d'IA, és a dir, no crea una forma per la qual les persones que es vegin perjudicades per un sistema d'IA puguin recórrer a l'AESIA per exigir protecció per un perjudici causat. En aquest sentit, sembla una autoritat creada bàsicament per gestionar els riscos del producte. La qüestió és que els riscos dels sistemes de IA poden implicar un impacte en els drets fonamentals de les persones. Per tant, caldrà incorporar una perspectiva basada en els drets fonamentals – sobre la qual l'AEPD té experiència.

4.3. AEPD i la protecció envers la discriminació algorítmica

L'RGPD disposa que cada estat membre establirà que sigui responsabilitat d'una o varies autoritats públiques supervisar l'aplicació de l'RGPD¹⁹³⁷, que serà/n independent/s¹⁹³⁸ i que inclourà/n entre les seves funcions -amb rellevància per a la discriminació algorítmica- controlar l'aplicació de l'RGPD i fer-lo aplicar; sensibilitzar al públic i promoure la comprensió dels riscos, normes, garanties i drets del tractament de dades, i la

¹⁹³⁵ Vegeu, per exemple, el cas d'Amazon i el seu algoritme de contractació que discriminava per raó de sexe: ADAMS-PRASSL, J., BINNS, R. i KELLY-LYTH, A., “Directly Discriminatory Algorithms”, *op. cit.*, 2023, pàg. 155 i 156.

¹⁹³⁶ Article 71 de la PRIA. També article 10.1.k del Real Decreto 729/2023, de 22 de agosto, por el que se aprueba el Estatuto de la Agencia Española de Supervisión de Inteligencia Artificial.

¹⁹³⁷ Article 51 de l'RGPD.

¹⁹³⁸ Article 52 de l'RGPD.

sensibilització de les persones responsables i encarregades del tractament de dades sobre les obligacions de l'RGPD; tractar les reclamacions presentades i investigar el motiu de la reclamació informant la persona reclamant del curs de la investigació, dur a terme investigacions sobre l'RGPD; promoure l'elaboració de codis de conducta; i fomentar la creació de mecanismes de certificació de protecció de dades¹⁹³⁹. Ha/n de tenir poders d'investigació (com ordenar al responsable o encarregat de tractament que faciliti qualsevol informació¹⁹⁴⁰), coercitius (com sancionar a la persona responsable o encarregada del tractament quan hagi infringit l'RGPD¹⁹⁴¹) i d'autorització i consultius (com acreditar els organismes de certificació¹⁹⁴²). Cal remarcar que l'RGPD també disposa que l'estat membre podrà establir per llei que l'autoritat de control tingui altres poders a més dels indicats¹⁹⁴³, com iniciar o exercitar accions judicials¹⁹⁴⁴.

A Espanya, l'LOPDGDD estableix que l'AEPD és una autoritat administrativa independent amb personalitat jurídica i plena capacitat pública i privada que actua amb plena independència dels poders públics en l'exercici de les seves funcions¹⁹⁴⁵ -sense adscriure's a cap ministeri, de la mateixa manera que l'Autoritat Independent, però no de l'AESIA, que s'integra dins el Ministeri d'Assumptes Econòmics i Transformació Digital¹⁹⁴⁶-, i elabora el seu propi pressupost¹⁹⁴⁷. L'article 47 de la LOPDGDD remet a l'article 57 de l'RGPD per establir les funcions de l'AEPD, de manera que no n'hi dona més de les que s'estipulen a l'RGPD, malgrat poder fer-ho¹⁹⁴⁸. S'encarrega de vetllar per la privacitat i la protecció de dades de la ciutadania, fomentar que les persones coneguin els seus drets i que les persones obligades coneguin la normativa. L'article 52 de l'LOPDGDD estableix un deure de col·laboració i obliga administracions públiques i

¹⁹³⁹ Article 57 de l'RGPD.

¹⁹⁴⁰ Article 58.1.a de l'RGPD. Per una anàlisi dels procediments tramitats per l'AEPD, vegeu: RUBIO RIVAS, E., "Régimen jurídico de los procedimientos tramitados por la Agencia Española de Protección de Datos en caso de posible vulneración de la normativa de protección de datos personales (Comentario al artículo 63 LOPDGDD)", a TRONCOSO REIGADA, A., (dir.) i VILLALBA CANO, L., (coord.), *Comentario al Reglamento General de Protección de Datos y a la Ley Orgánica de Protección de Datos Personales y Garantía de los Derechos Digitales*, Tomo II, Pamplona, Aranzadi, 2021, pàg. 2699-2729.

¹⁹⁴¹ Article 58.2.b de l'RGPD.

¹⁹⁴² Article 58.3.e de l'RGPD.

¹⁹⁴³ Article 58.4 de l'RGPD.

¹⁹⁴⁴ L'article 58.5 de l'RGPD disposa que "*Cada Estado miembro dispondrá por ley que su autoridad de control esté facultada para poner en conocimiento de las autoridades judiciales las infracciones del presente Reglamento y, si procede, para iniciar o ejercitar de otro modo acciones judiciales, con el fin de hacer cumplir lo dispuesto en el mismo*".

¹⁹⁴⁵ Article 44 de l'LOPDGDD.

¹⁹⁴⁶ Article 1.1 del Real Decreto 729/2023, de 22 de agosto, por el que se aprueba el Estatuto de la Agencia Española de Supervisión de Inteligencia Artificial.

¹⁹⁴⁷ Article 46 de l'LOPDGDD.

¹⁹⁴⁸ Article 47 de l'LOPDGDD.

particulars a proporcionar a l'AEPD les dades, informes, antecedents i justificants necessaris per dur a terme les seves tasques d'investigació. L'omissió d'aquest deure de col·laboració, està considerada una infracció molt greu¹⁹⁴⁹ que pot ser sancionada per la pròpia AEPD¹⁹⁵⁰.

En relació amb la discriminació algorítmica, per una banda, com a mecanismes preventius de rendició de comptes, l'AEPD s'encarrega de supervisar que es duguin a terme les obligacions en compliment del principi de responsabilitat proactiva que inclouen les avaluacions d'impacte i els requisits des del disseny (ja ens hem referit extensament en l'apartat 4 del capítol V del treball sobre la necessitat d'adaptar aquestes obligacions per tal que puguin ser útils en la detecció i mitigació de la discriminació algorítmica). Per altra banda, interessa la funció de l'AEPD d'ajudar a exercir els drets que confereix l'RGPD (per exemple, el dret d'oposició, d'informació o a no ser objecte de decisions individuals automatitzades), que poden tenir incidència en la detecció de casos de discriminació algorítmica. També la facultat que se li ha atorgat per obligar administracions públiques i particulars a col·laborar en les seves tasques d'investigació sota amenaça de sanció¹⁹⁵¹, lligada a la facultat de l'Agència per prendre decisions vinculants i imposar sancions¹⁹⁵², que poden suposar una ajuda enfront la discriminació algorítmica si un tractament de dades personals no respecta els principis i garanties establerts en l'RGPD, vulnera el principi de no-discriminació, i a través de la sanció de l'AEPD es finalitza el tractament que causa discriminació algorítmica. Cal tenir en compte però, que els sistemes d'IA poden haver-se desenvolupat sense emprar dades personals, i per tant, fora de la competència de les autoritats de protecció de dades. Serà quan el tractament de dades personals no respecti els principis establerts en l'RGPD que l'AEPD podrà actuar.

La pregunta en relació amb l'RGPD sobre quin grau de sensibilitat tenen les eines que estableix (avaluacions d'impacte i requisits des del disseny) per detectar formes de discriminació dinàmiques¹⁹⁵³, també sorgeix en relació amb la funció de l'AEPD. És a dir, la funció principal de l'AEPD és vetllar per la protecció de dades personals, no vetllar per tal que no es produeixi discriminació algorítmica. Això no obsta que l'AEPD

¹⁹⁴⁹ Article 72.ñ de l'LOPDGDD.

¹⁹⁵⁰ Article 68 de l'LOPDGDD.

¹⁹⁵¹ Article 52 de l'LOPDGDD.

¹⁹⁵² Article 47 de l'LOPDGDD.

¹⁹⁵³ HAKKARAINEN, J., "Naming something collective...", *op. cit.*, 2021, pàg. 13 i 14.

col·lateralment pugui contribuir a lluitar enfront la discriminació algorítmica, per exemple, quan aquesta vingui ocasionada per la violació dels principis de tractament de dades personals i amb la sanció per aquest tractament il·lícit també es previngui que s'ocasioni discriminació algorítmica. De la mateixa manera, la discriminació algorítmica també pot ser detectada a través de les eines que exigeixen responsabilitat proactiva, com les avaluacions d'impacte -per exemple, si a més d'una afectació a les dades personals es produeix discriminació algorítmica-, o previnguda a través dels requisits des del disseny.

Des de la perspectiva de Cotino Hueso, les autoritats independents properes al sector de les noves tecnologies -en particular les autoritats de protecció de dades- han de jugar un paper molt important -sobretot en la seva expertesa com a agències garants dels drets fonamentals. Per tenir aquesta rellevància haurien d'ampliar el seu espectre a altres drets i garanties constitucionals en joc¹⁹⁵⁴. Mantelero exemplifica aquesta necessitat d'ampliació de l'espectre de drets protegits en relació amb les avaluacions d'impacte de protecció de dades. L'autor, davant l'ús d'algoritmes en les tècniques modernes de processament de dades, suggereix que s'ha d'adoptar un espectre més ampli de les avaluacions d'impacte de protecció de dades, obligant les persones responsables de tractament a anar més enllà de la seva perspectiva tradicional de qualitat de dades i seguretat i considerar l'impacte del tractament de dades en els drets fonamentals i en els valors ètics i socials¹⁹⁵⁵. De la mateixa manera, les autoritats de protecció de dades haurien d'anar més enllà de la seva perspectiva tradicional centrada en la protecció de dades per considerar com el tractament de dades afecta els drets fonamentals de les persones i els valors ètics i socials.

Per una banda, des de la nostra perspectiva, les autoritats de protecció de dades es veuen limitades per protegir les persones de la discriminació algorítmica i altres drets fonamentals més enllà de la protecció de dades, malgrat que col·lateralment a la seva funció principal, en certes ocasions, ho puguin fer gràcies a l'accés a la informació que tenen i gràcies a la seva expertesa en treballar des de la perspectiva de protecció dels drets fonamentals. Per l'altra banda, com hem esmentat, Cotino Hueso i Mantelero plantegen la possibilitat que les autoritats de protecció de dades ampliïn el seu espectre a altres drets

¹⁹⁵⁴ COTINO HUESO, L., "Riesgos e impactos del big data, la inteligencia artificial y la robótica. enfoques, modelos y principios de la respuesta del derecho", *Revista General de Derecho Administrativo*, núm. 50, 2019, pàg. 33.

¹⁹⁵⁵ MANTELERO, A., "AI and Big Data: A Blueprint for a Human Rights...", *op. cit.*, 2018, pàg. 758.

i puguin assumir també, a part de la protecció de dades, la supervisió de com el tractament de dades afecta altres drets fonamentals de les persones. Està per veure si, atès el canvi de paradigma que suposa l'ús de sistemes d'intel·ligència artificial i la posició d'especial importància que té el tractament de dades en el funcionament dels sistemes algorítmics, les autoritats de protecció de dades ampliaran el seu espectre de protecció o no.

En qualsevol cas, és clar que les autoritats de protecció de dades tindran un rol de rellevància atès que el tractament de dades és una part molt important del funcionament d'un sistema algorítmic, malgrat en molts casos les dades que s'utilitzin per desenvolupar l'algoritme no siguin personals i no s'apliqui la protecció de dades. Per tant, i tal com es configura actualment l'esquema d'autoritats que poden tenir rellevància en la detecció, control i supervisió dels sistemes d'IA que poden ocasionar discriminació algorítmica, estudiem les possibilitats que obre per mitigar els efectes de la discriminació algorítmica la complementarietat, coordinació i cooperació entre l'AESIA, l'AEPD i l'Autoritat Independent.

4.4. Complementarietat, coordinació i cooperació entre l'AESIA, l'AEPD i l'Autoritat Independent

Cotino Hueso planteja que per l'especialització i caràcter tècnic de la IA, la seva impredictibilitat, la necessitat d'avaluació continua, la flexibilitat i adaptació als canvis i la necessitat d'especialitzar tècnics, interessats i especialistes, un element essencial de la regulació serà determinar qui fixa criteris, controla i supervisa l'ús de la IA¹⁹⁵⁶. Des de la nostra perspectiva, malgrat una autoritat concreta adopti el rol principal de fixació de criteris, control i supervisió, totes les altres autoritats que també es consideri que en l'àmbit de les seves competències s'empren sistemes de IA que tenen un impacte en els drets fonamentals (p. e. autoritats de control dels medicaments) o es consideri que el dret fonamental que tenen l'objectiu de protegir es veu afectat per sistemes de IA (p. e. autoritats de protecció de dades o d'igualtat i no discriminació), hauran de tenir eines per fixar criteris propis, per desafiar els sistemes d'IA que produeixin la vulneració de drets i

¹⁹⁵⁶ COTINO HUESO, L., "Riesgos e impactos del big data, la inteligencia artificial...", *op. cit.*, 2019, pàg. 31.

per establir una rendició de comptes significativa a la pràctica per a sistemes algorítmics discriminatoris¹⁹⁵⁷.

En aquest sentit, a més d'una autoritat de referència en matèria de IA, podria ser que en els pròxims anys es creessin, convergissin o s'especialitzessin diferents autoritats o organitzacions, que es redefineixin i es nodreixin de competències en matèria de IA i anàlisi massiva de dades¹⁹⁵⁸. Des d'aquesta punt de vista, en aquest apartat defensem que les tres autoritats analitzades anteriorment han de dur a terme una tasca complementària, coordinada i cooperant, per poder actuar enfront la discriminació algorítmica.

Les autoritats estudiades han de dur a terme una tasca complementària ja que actualment cap d'elles té les eines i les capacitats per assumir individualment el control de la discriminació algorítmica. L'Autoritat Independent està especialitzada en l'àmbit de la discriminació, l'AESIA estarà especialitzada en l'àmbit de la intel·ligència artificial i l'AEPD compta amb l'experiència d'integrar la perspectiva de la protecció dels drets fonamentals en un context tècnic i de gestió de riscos com és el tractament de dades. Això no obsta que, en un futur, la Autoritat Independent pugui especialitzar-se en l'àmbit de la discriminació algorítmica, com s'ha especulat que podia succeir si, en un futur, autoritats especialitzades redefinissin les seves competències en matèria de IA i anàlisi massiva de dades¹⁹⁵⁹.

A més, davant l'ús de sistemes algorítmics, poden dur a terme tasques diferents que acabin tenint repercussió en la lluita envers la discriminació algorítmica¹⁹⁶⁰. L'AESIA està facultada en la PRIA per supervisar tant la no introducció en el mercat de sistemes prohibits, com per revisar la documentació tècnica de sistemes de risc alt, el tractament de categories especials de dades amb la finalitat de garantir la vigilància, detecció i correcció de biaixos i els sistemes d'IA que en els espais de prova mostren resultats discriminatoris (cal tenir en compte que l'AESIA també estarà facultada per gestionar les queixes presentades per part de les persones perjudicades per un sistema d'IA si la proposta d'esmena del Parlament Europeu al text de la Comissió Europea es recull en el

¹⁹⁵⁷ WIERINGA, M., ““Hey SyRI, tell me about algorithmic accountability”...”, *op. cit.*, 2023, pàg. 4

¹⁹⁵⁸ COTINO HUESO, L., “Riesgos e impactos del big data, la inteligencia artificial...”, *op. cit.*, 2019, pàg. 33.

¹⁹⁵⁹ *Ídem, ibidem.*

¹⁹⁶⁰ COTINO HUESO, L., “Discriminación, sesgos e igualdad de la inteligencia artificial...”, *op. cit.*, 2023, pàg. 337. L'autor es refereix a específicament a que les autoritats de protecció de dades, supervisió dels algorismes i d'igualtat podran actuar potencialment quan es produeixi un incompliment normatiu.

text final¹⁹⁶¹). L'AEPD està facultada per l'RGPD i l'LOPDGDD per supervisar el compliment de les obligacions de realització d'avaluacions d'impacte de protecció de dades, la concreció de requisits des del disseny, la correcta implementació dels principis de responsabilitat proactiva i també pot participar en els espais de prova dels sistemes d'IA que mostrin violacions del tractament de dades i col·lateralment puguin causar discriminació algorítmica. De la mateixa manera, l'Autoritat independent podrà, entre d'altres, participar en els espais de prova instaurats per la PRIA per detectar discriminació algorítmica o confeccionar estadístiques que serveixin com a element de comparació en plets de discriminació. La funció, tant de l'AESIA com de l'AEPD, es concentra majoritàriament en què els sistemes d'IA compleixin els requisits exigits per la normativa (l'RGPD i la PRIA). Aquesta supervisió pot contribuir tant a la detecció de la discriminació algorítmica com a la no introducció en el mercat o la retirada del mercat d'un sistema d'IA que en generi.

També, com a control *ex post*, cada organisme actuant d'acord amb les seves potestats, podrà participar a que cessi l'ús d'un sistema d'IA que causa discriminació algorítmica. Per exemple, l'AESIA podrà sancionar i obligar a retirar del mercat intern un sistema algorítmic que no compleix els requisits establerts en Reglament de la IA; l'AEPD, a través de les seves potestats d'investigació, si detecta un tractament de dades personals que viola els principis de l'RGPD, podrà imposar una sanció, beneficiant col·lateralment la lluita envers la discriminació algorítmica, si aquest tractament n'estava produint i; de la mateixa manera, l'Autoritat Independent, a través de les seves potestats d'investigació, podrà detectar discriminació algorítmica i o bé interessar la seva sanció, o bé iniciar una acció de cessació per tal d'aturar l'ús del sistema algorítmic discriminatori.

A més, establir un marc de coordinació és d'interès per tal que actuïn harmònicament per a establir una acció comuna, essent aquesta la protecció de les persones enfront la discriminació que poden causar els sistemes algorítmics. En aquest sentit, alguns mecanismes són l'intercanvi d'informació entre les autoritats (per exemple, que existeixi un comitè específic en relació amb la discriminació algorítmica, que es reuneixi periòdicament, amb integrants de les tres autoritats on es comparteixi informació), l'intercanvi sistemàtic de memòries i informes, el desenvolupament de protocols conjunts

¹⁹⁶¹ EUROPEAN PARLIAMENT, *Amendments adopted by the European Parliament on 14 June 2023...*, *op. cit.*, 14 June 2023.

en casos de detecció de discriminació algorítmica, la compartició de comunicacions escrites o orals sobre resultats i mètodes de recerca, la creació de grups de treball per casos específics (per exemple, per llançar campanyes de prevenció envers la discriminació algorítmica o per fer seguiment de riscos) o l'establiment de guies per a la cooperació i actuació enfront la discriminació algorítmica¹⁹⁶².

Per exemple, el fet que l'organisme pugui rebre queixes individuals de les persones possibilita que les pugui analitzar i identificar si n'hi ha diverses similars, establint la concurrència d'un problema sistèmic. En aquest sentit, tant l'AEPD (si rep moltes queixes sobre tractament de dades personals il·lícit que col·lateralment provoqui discriminació algorítmica) com l'Autoritat Independent (si rep moltes queixes de discriminació que puguin ser identificades amb un mateix sistema algorítmic) poden detectar problemes sistèmics de discriminació algorítmica, coordinant-se per posar en comú informació i/o notificar-se entre elles la detecció d'un problema.

Finalment, han d'actuar dins un marc de cooperació, ja que malgrat l'autoritat especialitzada en discriminació sigui l'Autoritat Independent, l'AESIA i l'AEPD també han de prendre part per fer front a la discriminació algorítmica, atès que l'Autoritat Independent no tindrà les eines per poder assumir tot el pes de la supervisió i el control. Per exemple, l'Autoritat Independent és la única que té la perspectiva específica de la discriminació però necessitarà l'ús dels mitjans tècnics i l'especialització d'altres agències i autoritats, com l'AESIA, de manera que serà essencial que s'estableixin sistemes de cooperació. D'aquesta manera, per exemple, si presumptivament ha ocorregut una discriminació, la cooperació entre l'Autoritat Independent i l'AESIA, pot facilitar l'accés a informació relativa a la precisió del sistema d'IA sobre determinats grups de població per tal d'interrelacionar persones que s'han vist perjudicades i iniciar accions de cessació de l'ús del sistema d'IA en qüestió. Tanmateix, una altra perspectiva que s'ha plantejat és reconèixer a totes les autoritats que en el seu àmbit d'actuació convisquin amb l'ús de sistemes de IA, facultats àmplies d'accés i coneixement dels algoritmes, les lògiques i les grans bases de dades que es gestionin tant en el sector privat com en el

¹⁹⁶²Les idees plantejades en aquest apartat estan inspirades en el recull de tècniques de cooperació entre autoritats supervidores del mercat financer establertes en el següent article: KIM, H., "Cooperation and Coordination between the Financial Authorities: A Review of the Experiences of the United Kingdom, Norway, Sweden, and Korea", *Seoul Journal of Economics*, vol. 22, núm. 3, 2009, pàg. 415-429.

públic¹⁹⁶³. La situació òptima seria que totes les autoritats comptessin amb aquestes facultats i els recursos necessaris per fer front als reptes que planteja l'ús de sistemes d'IA, però en el cas que no fos així, la cooperació pot pal·liar l'aïllament en que es pot trobar l'Autoritat Independent davant la vulneració del dret a la no discriminació mitjançant l'ús de sistemes d'IA.

Ja existeixen esquemes de coordinació i cooperació entre agències reguladores per abordar l'àmbit de la IA. Per exemple, al Regne Unit l'any 2020 el Fòrum de Cooperació de Reguladors Digitals¹⁹⁶⁴ (DRCF) -integrat per l'Autoritat de competició i mercats¹⁹⁶⁵ (CMA), l'Autoritat nacional de protecció de dades (ICO), l'Oficina de Comunicacions¹⁹⁶⁶ (Ofcom) i l'Autoritat de Conductes Financeres¹⁹⁶⁷ (FCA)- va ser designat per promoure la col·laboració formal i la cooperació entre les autoritats reguladores en l'àmbit de la IA¹⁹⁶⁸.

En definitiva, és evident que per poder proporcionar una protecció col·lectiva-institucional a les persones envers la discriminació algorítmica -de la mateixa manera que envers totes les vulneracions que pugui ocasionar l'ús de sistemes d'IA- serà necessari, no només que es compti amb els recursos i la voluntat institucional, sinó que també que es concretin objectius comuns entre autoritats supervisores (tant a escala nacional com a escala europea) per poder brindar a les persones la possibilitat de desafiar les decisions que es prenguin sobre elles mitjançant sistemes algorítmics.

5. Conclusions parcials

En aquest capítol ens proposàvem respondre a la segona part de la cinquena qüestió de recerca del treball, és a dir, si cal una institució pública que coordini la lluita envers la discriminació algorítmica. La cinquena qüestió de recerca, que hem abordat en aquesta

¹⁹⁶³ COTINO HUESO, L., “Riesgos e impactos del big data, la inteligencia artificial...”, *op. cit.*, 2019, pàg. 33.

¹⁹⁶⁴ En anglès, Digital Regulators Cooperation Forum.

¹⁹⁶⁵ En anglès, Competition and Markets Authority.

¹⁹⁶⁶ En anglès, Office for Communications.

¹⁹⁶⁷ En anglès, Financial Conduct Authority.

¹⁹⁶⁸ ROBERTS, H. *et al.*, “Artificial intelligence regulation in the United Kingdom: a path to good governance and global leadership?”, *Internet Policy Review*, vol. 12, núm. 2, 2023, pàg. 5. (1-31)

tercera part del treball, pretenia respondre si una protecció col·lectiva-institucional envers la discriminació algorítmica podria protegir millor els col·lectius perjudicats per discriminació algorítmica, i alhora proporcionar eines suficients a les persones que s'han vist discriminades per poder desafiar les decisions discriminatòries preses mitjançant sistemes algorítmics.

Hem estudiat, doncs, els organismes d'igualtat nacionals, que sota la nostra perspectiva són els millor posicionats per adoptar un rol proactiu en defensa de les persones que es vegin perjudicades mitjançant sistemes algorítmics. La PRIA, com hem puntualitzat, no habilita cap acció col·lectiva en defensa dels interessos de persones que s'han vist perjudicades per una infracció d'un sistema algorítmic, i en el cas de la normativa antidiscriminació i l'RGPD sembla molt difícil aconseguir mitjans de prova que proveïxin elements de comparació necessaris per establir casos de discriminació. Per tant, el sistema d'obtenció de prova i tutela de les persones envers la discriminació algorítmica que veiem més factible, inclou que les persones -davant la intuïció d'haver patit discriminació per part d'un sistema algorítmic, ja sigui a través dels resultats que han obtingut per part del sistema algorítmic (seria interessant, com hem esmentat, que els sistemes algorítmics tinguessin un nom identificador i fos públic per a quina presa de decisió s'estan emprant), ja sigui a través de la informació obtinguda exercint els drets individuals que els confereix l'RGPD-:

1. Facin un queixa a un organisme d'igualtat nacional, en el cas espanyol a l'Autoritat Independent. També es podria donar el cas que l'organisme d'igualtat tingués notícia que s'està produint discriminació algorítmica per qualsevol mitjà¹⁹⁶⁹ (arran de la informació obtinguda per part de l'organisme pròpiament o de la informació que pot obtenir a través de la documentació tècnica de sistemes d'IA de risc alt) i iniciés una investigació.
2. L'organisme d'igualtat nacional, obtingués, en el cas que el sistema que presumptivament està produint discriminació fos classificat de risc alt pel Reglament de la IA, la documentació tècnica presentada per la persona proveïdora

¹⁹⁶⁹ Per exemple a través d'un sistema anònim de denúncia d'irregularitats (en anglès s'anomena whistleblowing).

del sistema o, emprés les prerrogatives d'investigació de les quals disposa per obtenir informació que esclarís si s'ha produït discriminació algorítmica.

3. A partir d'aquesta informació tècnica i amb la informació proporcionada a través de les queixes o a través de les seves facultats d'investigació, l'organisme nacional d'igualtat o bé actués com a òrgan quasi-jurisdiccional i prenguéss una decisió sobre el sistema algorítmic discriminatori i sancionés, si fos pertinent, l'empresa o administració pública que estigués emprant el sistema algorítmic discriminatori; o bé recollís indicis clars que les persones pertanyents a una (o diverses) categoria(es) protegida(es) per la legislació han patit discriminació *prima facie* causada per l'ús del sistema algorítmic i,
4. O bé iniciés d'ofici accions en defensa de l'interès general o d'interessos difusos per aconseguir el cessament de l'ús del sistema algorítmic que està causant discriminació (podent emprar estadístiques confeccionades pel propi organisme com element comparador o d'altres a les quals tingués accés); o bé iniciés accions de cessació o participés com a part en accions en defensa d'interessos col·lectius -si les persones perjudicades estiguessin degudament identificades-; o bé proporcionés informació a les persones perjudicades per tal que poguessin dur a terme accions en defensa dels seus interessos o d'interessos (pluri)individuals homogenis, podent, les persones individuals implicades sol·licitar també reparació pel dany moral causat.

Aquest rol proactiu per part de l'organisme nacional d'igualtat per tal de proporcionar una protecció col·lectiva-institucional amb efectivitat, s'hauria de complementar amb l'administració d'una plataforma pública que facilités la gestió de queixes per discriminació algorítmica, interrelacionés persones perjudicades i pogués detectar problemàtiques sistèmiques per iniciar litigació estratègica i iniciar accions de cessació amb l'objectiu de fer cessar l'ús de sistemes algorítmics discriminatoris. També l'organisme hauria d'estudiar els casos per abordar de manera unificada la discriminació interseccional, la participació en els espais de prova dels sistemes algorítmics que preveu la PRIA, el manteniment d'una llista de sistemes algorítmics que s'haguessin provat discriminatoris, i la coordinació amb les agències de supervisió de la intel·ligència artificial i de protecció de dades.

A Espanya, l'Autoritat Independent pot dur a terme aquestes funcions malgrat que n'hi ha certes que no estan clarament especificades en la legislació, ja que no compta amb capacitat per funcionar com a òrgan quasi-jurisdiccional i és només quan protegeix interessos difusos que pot iniciar accions en el seu propi nom sense necessitar l'autorització de la víctima. Tanmateix, ha d'elaborar informes i estadístiques de caràcter periòdic, pot iniciar d'ofici o a instància de tercers investigacions sobre l'existència de possibles situacions de discriminació que plantegin especial gravetat o rellevància, i les administracions públiques i particulars han de prestar col·laboració necessària, possibilitant l'accés a informació i dades que l'Autoritat Independent necessiti en les seves investigacions. Considerem, doncs, que no obstant que l'Autoritat Independent disposa de l'especialització i les funcions necessàries per posicionar-se com a òrgan de referència en la lluita envers la discriminació algorítmica, serà d'especial rellevància establir marcs de coordinació i cooperació amb altres autoritats de referència quan es tracta de la supervisió de l'ús de sistemes d'IA, com l'AESIA o l'AEPD, ja que no només poden ser d'ajuda, per la seva especialització en control i supervisió de la IA o d'integrar la perspectiva de la protecció dels drets fonamentals en un context tècnic i de gestió de riscos respectivament, sinó que inclús poden ser complementàries a l'actuació de l'Autoritat Independent.

En conseqüència, concloem que és necessària una institució pública que coordini la lluita envers la discriminació algorítmica. La instauració d'un sistema de protecció col·lectiva-institucional envers la discriminació algorítmica millora la situació de la persona perjudicada principalment perquè li dona la possibilitat de recórrer a un organisme amb més recursos (p. e. possibilitat d'accedir a informació i proves de discriminació, o a estadístiques que poden funcionar com a element comparador), que a la seva vegada també pot actuar d'ofici, i que té com a objectiu la protecció del col·lectiu perjudicat, de manera que l'abast de les seves accions pot ser molt més ampli.

CONCLUSIONS

Hem estructurat aquest treball en vuit capítols que ens han permès respondre a les preguntes de recerca que ens plantejàvem a la introducció. Al final de cada capítol es respon a la pregunta de recerca corresponent. Hi ha una línia comuna en l'estudi consistent en la recerca d'eines efectives per tal que les persones discriminades per la presa de decisions mitjançant sistemes algorítmics puguin desafiar-les, amb especial atenció a les eines que contribueixin a la protecció del col·lectiu perjudicat per la discriminació. A través d'aquesta recerca podem plantejar diverses conclusions generals (1) i formular diverses propostes concretes per millorar la protecció de les persones discriminades (2). Cal puntualitzar que hi haurà novetats legislatives que s'incorporaran a mida que siguin adoptades en una versió posterior del treball de cara a la seva publicació.

1. Conclusions generals

1.1. Els sistemes de presa de decisió mitjançant algoritmes aplicats a contextos socials són eines que amplifiquen la discriminació ja existent en la societat. Per com operen, es trasllueix que assumeixen perfectament els sistemes d'opressió/subordinació existents, els sistematitzen i els exploten tot beneficiant-se de l'absència d'una normativa capaç de desafiar-los.

Una de les causes és el canvi de paradigma que s'ha produït en l'explotació de la informació sobre les persones. Els sistemes algorítmics funcionen traduint conceptes complexos en dades amb la finalitat de ser més eficaços en la presa de decisions (però, no obstant, simplificant la realitat). Abans de l'anàlisi massiva de dades, les decisions eren preses per una o diverses persones d'acord amb els propis prejudicis i estereotips (conscientment o inconscientment). En canvi, una decisió presa mitjançant l'anàlisi massiva de dades es pren d'acord amb l'assumpció dels sistemes d'opressió i subordinació dominants en una societat, ja que simplifica la informació personal i margina aquelles persones que no realitzen el mateix comportament que la majoria o de

les quals no disposa de tanta informació.

A més, les decisions preses mitjançant sistemes algorítmics dificulten que les persones detectin que estan essent subjectes a discriminació i destrueix tota concepció de grup. És a dir, malgrat que la discriminació algorítmica es produeixi arran de la reproducció de sistemes d'opressió envers col·lectius determinats de població, les persones amb les mateixes característiques que es veuen discriminades no desenvolupen cap consciència de grup, cap "nosaltres", generant que no es consolidi cap grup subjecte d'una acció comuna capaç d'atacar les relacions existents de dominació. Només existeixen subjectes discriminats isolats incapaços de conformar un poder contrari i de qüestionar realment la discriminació a la qual han estat subjectes¹⁹⁷⁰.

Els sistemes algorítmics aplicats a contextos socials participen a gran escala a classificar persones segons uns criteris preestablerts per grups de població privilegiats, determinant les seves oportunitats. Participen a que els grups socials estiguin cada cop més delimitats i que aquells que en l'estructura social tenen una posició subordinada, tinguin menys possibilitats entre les quals escollir. La discriminació algorítmica -igual que la tradicional, però en el cas de la primera, de manera més evident- és la resposta als prejudicis i estereotips sobre col·lectius determinats de persones adquirits pel sistema algorítmic.

Atesa la rapidesa dels avenços tecnològics, la protecció dels drets vulnerats sempre anirà molt per darrera de la tecnologia. Malgrat això, la immensitat del repte no pot privar les persones de formular propostes per abordar normativament aquest fenomen.

1.2. El dret antidiscriminatori europeu presenta una estructura i concepció teòrica que dificulta l'abordatge de la discriminació produïda mitjançant sistemes algorítmics. La manca de perspectiva interseccional dificulta la lluita envers els sistemes d'opressió que operen a la societat, i per consegüent, la lluita envers la discriminació algorítmica.

La normativa antidiscriminació europea està fragmentada segons la característica protegida i l'àmbit d'aplicació material. És una normativa marcadament orientada per criteris anticlassificatoris, i el desenvolupament de la normativa s'ha fet a través de

¹⁹⁷⁰ Per una explicació sobre la impossibilitat en un examen digital de desenvolupar concepció de grup, vegeu: HAN, B-C, *En el enjambre*, Barcelona, Herder, 2021, pàg. 25-32.

directives, que deixen ampli marge d'apreciació als estats membres, causant diferències en la protecció. La configuració de la normativa es basa en una concepció de la discriminació que obvia l'existència de sistemes d'opressió que s'interrelacionen entre ells i inclús, en ocasions, els reforça (p. e. al possibilitar l'apreciació de discriminació inversa), centrant-se en donar resposta a accions discriminatòries puntuals i individuals que, tot i facilitar l'aplicació judicial, negligeixen la creació de canvis de caràcter sistèmic. Partint de la base que les persones formem part dels sistemes d'opressió, que són múltiples i diferenciats -és a dir, malgrat que una persona sigui subordinada en un sistema d'opressió (dona), pot ser opressora en un altre (de classe)¹⁹⁷¹-, la normativa antidiscriminació no pot obviar que tots tenim prejudicis i que les decisions es prenen en un context fonamentalment desigual. S'ha plantejat la possibilitat d'instaurar, com a criteris auxiliars interpretatius per part dels tribunals que tracten casos de discriminació, la discriminació estructural (que pot ajudar a interpretar el tracte que produeixi una discriminació a la llum de la posició social estructural de subordinació i a entendre que els algoritmes repliquen i perpetuen les estructures d'opressió) i la discriminació interseccional (que pot ajudar a millorar la interpretació del resultat d'una decisió presa mitjançant un algoritme, que en molts casos, no basarà la discriminació en un atribut protegit, sinó que ho farà en base a múltiples atributs protegits, possibilitant que es tingui en compte que si determinades decisions automatitzades no comporten un tractament menys favorable en relació amb un atribut protegit, potser sí que el comporten en base a diversos atributs protegits combinats). No obstant això, cal ser conscients que la normativa antidiscriminació europea, tal com està plantejada fins ara, és un mecanisme inadequat per adreçar estructures d'opressió profundament arrelades socialment, i per tant, les eines que ofereix per adreçar-les també són limitades. Són aquestes eines les que hem pretès potenciar per adreçar la discriminació algorítmica en plantejar un sistema de protecció col·lectiva-institucional.

1.2.1. L'ús de classificacions raonables per prevenir la discriminació pot no ser útil en relació amb la discriminació algorítmica.

¹⁹⁷¹ Sobre la diversitat d'opressions socials que pot partir una persona, vegeu la carta titulada "D'una obrera a una senyora que es dedica a fer conferències parlant de l'emancipació de la dona" a: CAPMANY, M. A., *Cartes impertinents*, Barcelona, Horsori, 2009, pàg. 89-92.

Per instrumentalitzar tant el dret antidiscriminatori com el dret a la protecció de dades personals, actualment és essencial l'ús de classificacions raonables (atributs protegits en el cas de la normativa antidiscriminació i categories de dades sensibles en el cas de la normativa de protecció de dades). Tanmateix, aquestes classificacions raonables causen indefectiblement l'exclusió d'altres atributs o categories, l'ús dels quals, en el context dels sistemes algorítmics, poden reforçar les desigualtats socials. A més, en el cas de la protecció de dades, la prohibició de l'ús de les categories sensibles pot inclús dificultar la detecció de biaixos. El plantejament en l'article 10.5 de la PRIA, d'habilitar el tractament de categories especials de dades per garantir la vigilància, detecció i correcció dels biaixos associats al sistema d'IA de risc alt, sembla una bona proposta, sempre que s'estableixin salvaguardes reals i efectives i no s'acabi desprotegint als col·lectius més vulnerables.

Igualment, tal com plantejades les normatives, és necessari establir de manera consensual quins atributs personals o categories de dades han de rebre especial protecció, però cal tenir en compte que només entren en aquestes classificacions raonables aquelles que el context social i polític està disposat a admetre en una societat determinada, quan hi ha molts altres atributs i categories de dades que causen subordinació dins els sistemes d'opressió i que queden fora de la protecció legislativa. Per això sembla d'interès considerar els atributs i categories de dades que enumera la legislació com a exemples, però a la vegada incloure en la protecció altres atributs i categories sensibles que usats de manera anàloga també reforcin desigualtats socials (a proposta de la part demandant, ha de ser el tribunal qui decideixi si els atributs o categories són susceptibles de reflectir biaixos, prejudicis o estereotips de manera comparable). En aquest sentit, interessa puntualitzar que només és lògic considerar com a discriminatòries aquelles conductes que reforcen els sistemes d'opressió, no qualsevol tracte desigual.

1.2.2. La distinció entre discriminació directa i indirecta en el context de la discriminació algorítmica no és operativa

La jurisprudència del TJUE actualment segueix el següent esquema per distingir entre discriminació directa i discriminació indirecta: estem davant de discriminació directa si s'ha fet un ús exprés de la categoria protegida per la legislació (p. e. cas Mangold); si es pot provar que un criteri, malgrat ser formalment neutre, s'ha escollit per raons

relacionades amb una categoria protegida per la legislació (vegeu l'argumentació que fa el tribunal en el cas CHEZ), o si el criteri usat és inseparable o està inextricablement vinculat a una categoria protegida per la legislació (p. e. cas Dekker i Tadao Maruko). Si s'empra un criteri formalment neutre i no es troba en els casos anteriors (p. e. cas Jenkins i Bilka-Kaufhaus), estem davant de discriminació indirecta. Sembla llavors que la distinció entre la discriminació directa i la discriminació indirecta té una relació clara en destriar si la persona que pren la decisió era conscient o podria haver estat conscient de la relació entre l'atribut protegit i el resultat, o no. En el cas de sistemes algorítmics, es poden donar les següents situacions: si l'algoritme té directrius concretes per tractar desfavorablement un grup determinat de població que comparteix un atribut protegit, si el sistema algorítmic empra dades que venen donades en base a un patró històric de discriminació, si empra atributs equivalents a les categories protegides o que són inseparables o inextricablement vinculats a les categories protegides, estem davant de discriminació directa. Tanmateix, serà pràcticament impossible provar el coneixement per part de les persones que han emprat el sistema algorítmic de la relació o la implicació dels atributs protegits en l'obtenció d'un resultat discriminatori per part del sistema, de manera que atesa l'opacitat inherent al funcionament del sistema algorítmic, el més probable serà considerar que s'ha produït discriminació indirecta en vista del resultat discriminatori.

Hi ha llavors diverses qüestions sobre les quals parar atenció:

- Confiar que la persona que s'ha vist discriminada pugui demostrar -després d'haver demostrat que ha patit discriminació- que aquesta ha estat directa, ja que la persona que ha emprat el sistema de presa de decisió era conscient que els resultats serien discriminatoris o perquè s'han utilitzat atributs directament protegits o equivalents, inseparables o inextricablement vinculats, no sembla molt probable.
- Si no es pot demostrar aquesta relació, la discriminació causada per decisions preses mitjançant sistemes algorítmics recaurà en la doctrina de la discriminació indirecta, que permet justificacions objectives (concepte que també és molt discutible des de la perspectiva antisubordinatòria atesa la impossibilitat d'avaluar l'objectivitat en una societat fonamentalment desigual) i presenta dificultats de prova, sobretot en relació amb l'establiment tant d'un grup desavantajat com d'un grup comparador, ja que rarament es formarà un grup amb suficient consistència

com per qüestionar la decisió discriminatòria (com hem esmentat en el punt 1.1 de les conclusions) i els sistemes algorítmics poden excloure persones del resultat directament sense que elles en siguin conscients i discriminar de manera subtil i no evident en base a múltiples factors (atomitzant tant els grups perjudicats com els comparadors).

- Malgrat que, segons el TJUE, provar intencionalitat va deixar de ser rellevant per apreciar discriminació, sí que sembla seguir essent rellevant perquè el tribunal apreciï discriminació directa o indirecta. És a dir, en la lectura dels casos en què el TJUE dona criteris interpretatius per distingir la discriminació directa de la indirecta, trasllueix que es busca dilucidar si el criteri controvertit s'ha emprat específicament per discriminar (encara que s'hagi usat amb subterfugis -atribut indissociable o inextricablement vinculat-) o ha acabat discriminant però era difícilment probable que es pogués saber d'avançada. S'ha de tenir en compte que la distinció va ser pensada en un moment en què era factible (malgrat ésser complicat) provar que els resultats discriminatoris s'havien donat en un context en el qual qui prenia la decisió era conscient o podia ser conscient que l'acció, criteri o pràctica emprat generava resultats discriminatoris (discriminació directa, p. e. declaració pública en el cas Feryn, Accept o Rete Lenford) o no (discriminació indirecta). En canvi, en el cas de la discriminació algorítmica, aquesta distinció pot ser impossible de fer.
- Des de la perspectiva antisubordinatòria és una distinció poc rellevant. Està plantejada des de la perspectiva en què no existeixen sistemes d'opressió en els quals tots participem, sinó que hi ha certes persones que no actuen com seria desitjable i duen a terme accions discriminatòries (essent-ne conscients -discriminació directa-, o essent-ne inconscients -discriminació indirecta-). Tenint en compte que efectuar aquesta distinció quan estan implicats sistemes algorítmics és molt complicat, no sembla configurar un recurs lògic per enquadrar la discriminació algorítmica.
- Canviant de perspectiva, en estar demostrat que tots els algoritmes aplicats a contextos socials són discriminatoris, ja que estan creats en una societat desigual estructurada en base a sistemes d'opressió, s'hauria de canviar d'enfocament i normativitzar un cribratge de precisió exigible a cada sistema, aplicable a contextos

socials per qualsevol grup determinat de població (incloent la perspectiva interseccional). Llavors, l'apreciació de discriminació dependria de si el sistema demostra o no que assoleix els estàndards de precisió exigits (vegeu punt 2.1 de les conclusions).

1.2.3. Malgrat que les persones que pateixen discriminació algorítmica es trobin en una clara situació de desigualtat, la normativa antidiscriminació no dona eines per presentar indicis de discriminació.

Les persones que formen part de grups socialment vulnerables no només viuen en una posició subordinada dins els sistemes d'opressió que operen en la societat, sinó que quan es troben amb un conflicte puntual susceptible de desenvolupar una protecció jurídica concreta, la configuració del dret antidiscriminatori les posiciona en igualtat amb l'altra part pel que fa a l'obtenció d'indicis. De manera que no garanteix o no facilita l'obtenció d'indicis per part de la persona discriminada. L'única facilitat que s'ha desenvolupat arran de la jurisprudència del TJUE ha estat l'admissió que la denegació total de l'accés a la informació per la part demandada pugui constituir un dels factors que s'han de tenir en compte en l'establiment d'un cas *prima facie* (p. e. cas Meister). Que posteriorment canviï el repartiment de la càrrega de la prova és anecdòtic si és impossible plantejar indicis de discriminació. Si la discriminació es produeix mitjançant sistemes algorítmics, s'exacerba la dificultat d'accés a prova, atesa l'opacitat i subtilitat del sistema. Addicionalment, el més probable és que per plantejar indicis de discriminació sigui necessària informació sobre com el sistema algorítmic tracta a grups determinats de població, informació actualment no disponible per al públic general.

1.3. L'RGPD no és prou útil per proporcionar prova envers la discriminació algorítmica. Estableix una regulació molt feble de les decisions automatitzades, no aborda directament els biaixos algorítmics, estableix unes exigències d'informació i transparència difícilment operatives en un cas de discriminació algorítmica, i planteja una perspectiva marcadament individual i anticlassificatòria. Tanmateix, presenta elements que poden donar idees (que necessiten desenvolupament) per facilitar l'accés a la prova de discriminació algorítmica i detectar-la.

1.3.1. Les persones que es veuen discriminades mitjançant sistemes algorítmics no disposen de suficients eines de transparència com per ser-ne conscients. La informació que es dona a través de l'RGPD, per ser útil, ha de ser en forma de explicacions contrafàctiques.

Els articles 12.1, 13.2 (f), 14.2 (g) i 15.1 (h) i el considerant 71 de l'RGPD proporcionen a les persones que es veuen subjectes a una decisió únicament automatitzada la possibilitat d'obtenir informació sobre la presa de decisió, però la informació que es dona és a nivell individual i pot interferir amb els drets de propietat intel·lectual, pot ser tan complexa que no es pugui traduir en informació que sigui possible facilitar a una persona no especialitzada, pot ser intel·ligible o no ser d'ajuda per proporcionar proves en un cas de discriminació algorítmica. En aquest sentit, hem pres la proposta de Wachter, Mittelstadt i Russel¹⁹⁷² (que defensen que en comptes de centrar la normativa en explicar com les decisions s'han pres internament, cal centrar-la en obligar a donar informació sobre quins fets externs poden ser diferents per tal que els resultats de les decisions preses mitjançant sistemes algorítmics siguin favorables), i l'hem instrumentalitzat per poder generar informació útil davant la discriminació algorítmica. En el cas de la denegació d'un préstec, si una explicació contrafàctica donaria les raons per les quals s'ha generat una decisió particular (p. e. tenir un salari baix), proporcionaria les bases per contestar-la (p. e. la persona responsable del tractament ha emprat dades imprecises sobre el salari de la persona sol·licitant) i aportaria recomanacions sobre com rebre els resultats volguts en el futur (p. e. si es tingués un salari de 4.000 euros més a l'any, el resultat podria haver estat positiu); en casos en què l'algoritme hagi incorregut en discriminació en la denegació d'un préstec, la raó particular seria (p. e. ser dona), no hi hauria bases per contestar-la (ja que no hi ha dades imprecises sobre el fet de ser dona) i finalment, el resultat volgut en un futur perquè l'algoritme donés el resultat desitjat seria (p. e. no ser dona). Potser el cas de ser dona és molt evident, però podrien revelar, per exemple, que per obtenir una decisió favorable t'has de canviar de lloc de residència o de codi postal. Una decisió presa en base a aquests criteris seria susceptible de ser contestada. El que

¹⁹⁷² WACHTER, S., MITTELSTADT, B., i RUSSEL, C., "Counterfactual explanations without...", *op. cit.*, 2018, pàg. 841-887.

interessa saber, igual que en les decisions preses per éssers humans, seria quins criteris han fonamentat la decisió, per poder, en cas de disconformitat, rebatre'ls.

Inicialment les explicacions contrafàctiques es plantegen per tal que les persones puguin saber de quina manera poden canviar la decisió presa per l'algoritme i millorar la seva situació personal. En canvi, en la nostra proposta, les explicacions contrafàctiques ens semblen d'interès atès que creiem que proporcionen la informació pertinent per saber en quins factors es basa l'algoritme per prendre la decisió, i com aquests factors poden revelar fàcilment un tracte discriminatori. També, perquè incorporen dades de funcionament de l'algoritme (quins factors ha valorat l'algoritme en la presa de decisió) en una explicació individual.

Aquesta tipologia d'informació podria ser útil en situacions on fos evident que el criteri emprat té relació amb el resultat discriminatori, atès que amb la informació individual obtinguda la persona podria tenir suficients indicis com per presentar un cas *prima facie* (discriminació directa), però en els casos en què la relació entre criteri i resultat discriminatori no fos evident, malgrat que a nivell individual pogués donar-se certa informació, sempre seria necessària una altra perspectiva grupal relativa a l'impacte del criteri sobre un determinat grup de població (discriminació indirecta).

Problema: No obstant això, s'ha de tenir en compte tant la probable impossibilitat, quan es duu a terme la presa de decisions mitjançant sistemes algorítmics, per aclarir la multitud de factors que han contribuït a formar cada coeficient, com també que el llenguatge tècnic i la documentació existent siguin suficients per explicar els detalls necessaris a qualsevol persona que no tingui coneixements previs especialitzats.

1.3.2. La detecció de la discriminació algorítmica a través de les eines de l'RGPD sempre estarà supeditada a un tractament il·lícit de dades personals.

Aquest fet limita ostensiblement l'abast de l'RGPD com a eina útil envers la discriminació algorítmica. Tanmateix, de la protecció de dades des del disseny i per defecte, pot extreure's la idea d'igualtat des del disseny i per defecte com a estàndard en el desenvolupament dels algoritmes (per tal d'estructurar-los de manera que siguin sensibles a les formes predominants de discriminació i inclús corregeixin desequilibris

socials entre membres de grups avantatjats i desavantatjats). També, de les avaluacions d'impacte relatives a la protecció de dades, pot extreure's la idea d'avaluacions d'impacte algorítmiques (en el cas que es vegin implicats algoritmes en el tractament de dades), que puguin detectar col·lectius que es poden veure perjudicats pel tractament de dades generat pel sistema algorítmic. Igualment, remarquem, la detecció de discriminació algorítmica seria un èxit col·lateral dins l'objectiu de l'RGPD que és la protecció de dades personals.

1.4. La Proposta de Reglament de la IA presenta obligacions per als proveïdors de sistemes d'IA útils tant per la detecció com per al desafiament de la discriminació algorítmica. Ho fa a mode perifèric, atès que es presenta com una regulació de seguretat en el producte que té com a prioritat el bon funcionament del mercat intern -de manera que la protecció dels drets fonamentals, malgrat justificar la regulació, queda en segona pla. Nogensmenys, imposa l'obligació de crear una base de dades pública de sistemes d'intel·ligència artificial de risc alt (art. 60), estableix espais de prova on es podran detectar prèviament biaixos discriminatoris en el sistema d'IA (art. 53), possibilita el tractament de dades sensibles amb la finalitat de detectar discriminació (art. 10.5) i obliga els proveïdors dels sistemes de risc alt a crear documentació tècnica que pot ser molt rellevant en els casos de discriminació algorítmica, ja que ha de contenir la precisió del sistema en diferents grups de població (article 11 en relació amb els punts 2.g i 3 de l'annex IV). Tanmateix, només dona accés a la informació a organismes notificats, autoritats nacionals competents i, amb l'autorització d'aquestes, a organismes d'igualtat dels estats membres (article 64.3 i considerant 79 de la PRIA).

També, malgrat que prohibeix l'ús de certs sistemes d'IA (art. 5.1) que porten a un tracte perjudicial o desfavorable per a certes persones o grups, es fa en un context relativament restringit. Quant als sistemes de risc alt, reconèixer tots els efectes discriminatoris i danys potencials que poden patir grups oprimits arran de l'ús d'aquests sistemes i no prohibir-los, devalua la importància que la normativa dona a la discriminació.

1.5. La governança publicoprivada subjacent tant en l'RGPD com en la PRIA és una fórmula deficient per protegir les persones de la discriminació produïda arran de sistemes algorítmics, essent força inversemblant el plantejament que les entitats privades

s'esmercen en vetllar per l'interès públic. Cal ser conscients que un sistema regulador cooperatiu entre empresa privada i administració pública -que entenem necessari atès el desequilibri d'expertesa entre àmbit privat i àmbit públic- és deficient per la protecció de drets fonamentals. Floridi afirma que és hora de forçar les grans corporacions a acatar la legislació, substituint el dret tou per la llei¹⁹⁷³. Tanmateix, és evident que les reticències per part la UE d'imposar obligacions realment restrictives als sistemes d'IA provenen de la por que aquestes puguin causar una disminució de l'activitat econòmica. En una societat capitalista, a més, unes eines que prometen rapidesa, eficàcia i reducció de costos són impossibles de passar per alt o prohibir -malgrat tinguin efectes negatius pels drets de les persones. Davant d'aquesta conjuntura, aquest treball s'ha centrat en la proposta d'un sistema per incentivar i millorar les eines reactives, per tal que, si més no, les persones tinguin suport i eines per desafiar els sistemes que puguin causar-los perjudicis (vegeu punt 2.2 i 2.3 de les conclusions).

La justificació per instaurar una governança publicoprivada (incloure els recursos i l'expertesa del sector privat en la regulació pública) no compensa els riscos que genera (que el sector privat es converteixi *de facto* en autoregulat ja sigui perquè amb la seva expertesa i recursos pot definir i incidir discrecionalment en conceptes clau de la regulació o ja sigui perquè el poder públic no pot instaurar una rendició de comptes eficient).

La governança publicoprivada suposa diverses implicacions:

- Les eines de gestió del risc i de responsabilitat proactiva (avaluació d'impacte relativa a la protecció de dades, protecció de dades des del disseny i per defecte, i auditories) plantejades al llarg de l'RGPD són controlades per les mateixes empreses que les apliquen. Pel que fa a les avaluacions d'impacte i a les auditories, no hi ha obligació de revelar al públic general els resultats. De la mateixa manera, per mitigar els resultats discriminatoris de sistemes algorítmics que s'apliquen en contextos socials, la inclusió de criteris sobre la no discriminació en codis de conducta i com a requisits per obtenir una certificació, segell o marca pot ser una bona opció, malgrat que existeix el risc evident que esdevinguin processos merament formalistes i normatius sense més rellevància.

¹⁹⁷³ FLORIDI, L., "The End of an Era: from Self-Regulation to Hard Law for the Digital Industry", *Philosophy and Technology*, vol. 34, núm. 4, 2021, pàg. 621 i 622.

- L'habilitació del tractament de categories especials de dades per garantir la vigilància, detecció i correcció dels biaixos associats al sistema d'IA de risc alt que planteja l'article 10.5 de la PRIA requereix l'establiment de salvaguardes reals i efectives per tal que no s'acabi desprotegint als col·lectius més vulnerables, salvaguardes que han d'aplicar les mateixes empreses o organitzacions que duran a terme el tractament de dades.
- L'estructura de la PRIA basada en la implementació de certs requisits per a la comercialització de sistemes de risc alt, que en la majoria dels casos són autoavaluables pel mateix proveïdor que vol comercialitzar-los, no sembla una protecció suficient. A més, molts sistemes d'IA podran considerar-se de risc baix, i no generaran obligacions, quan potser aplicats a un context determinat tindran una incidència molt més negativa d'aquella esperada.
- Els sistemes d'IA comporten qüestions ètiques difícilment traduïbles a estàndards harmonitzats. Atès que les implicacions dels sistemes d'IA van més enllà de mers tecnicismes, els estàndards haurien de ser discutits democràticament i no creats per les organitzacions europees de normalització (organitzacions de dret privat). Tant pel marcatge "CE" de conformitat, com pel seguiment posterior a la comercialització, la figura del proveïdor del sistema d'IA ostenta molt poder.

En definitiva, cal tenir en compte, i sobretot quan es pretenen instrumentalitzar aquestes regulacions en benefici de la protecció de drets fonamentals, que tant en l'RGPD com en la futura regulació de la intel·ligència artificial, es pot advertir -més enllà de la finalitat protectora- l'objectiu de la UE de crear unes bases per aconseguir un mercat digital únic i evitar que es continuïn produint obstacles des dels estats membres per al mercat intern.

1.6. Que es produeixi discriminació algorítmica depèn, en part, del context en el qual el sistema es desenvolupa. La normativa tant en l'RGPD com en la PRIA no sembla focalitzar l'atenció en què els sistemes algorítmics evolucionin amb el temps i canviïn en relació amb l'avaluació, detecció i mitigació dels riscos inicials.

- Sembla difícil que les auditories estipulades en l'RGPD i les avaluacions de protecció de dades es puguin aplicar a sistemes que evolucionen amb el temps i

que puguin evitar que els paràmetres que les defineixen tinguin els mateixos biaixos (explícits i implícits) que es poden trobar en el tractament en si.

- Tant l'exigència d'exactitud com els articles 15, 16, 17, 18, 20 i 21 de l'RGPD, poden participar a millorar la qualitat de les dades que tenen una incidència directa en els biaixos algorítmics causants de la discriminació algorítmica. No obstant això, l'RGPD planteja la qualitat de dades de manera estàtica, és a dir, les dades que hi ha es poden canviar, fet que no s'adapta als sistemes algorítmics que evolucionen a través de la correlació de dades que no es podran alterar retrospectivament. A més, només genera eines per fer front a la vessant tècnica d'aquesta problemàtica, havent-hi altres accions necessàries per mitigar l'aparició de biaixos i reduir-ne l'impacte, com la sensibilització social envers els biaixos inconscients, equips de treball amb diversitat d'edat, ètnia, cultura o gènere, ús de dades com més representatives millor, promoció d'una IA diversa i inclusiva, comprensió i mesura de biaixos, generació de noves bases de dades d'entrenament o que les persones que creen bases de dades en defineixin les característiques i les limitacions per tal que quan s'utilitzin es tingui tota la informació pertinent per al seu ús.
- La PRIA, de manera predominant, imposa obligacions als proveïdors i no als usuaris dels sistemes d'IA, que en virtut de la seva responsabilitat pels riscos relatius a com el sistema està sent usat, també haurien de tenir obligacions en referència als drets fonamentals.
- En la PRIA la classificació del risc de cada sistema s'estableix d'avançada, fet problemàtic atès que en la discriminació algorítmica el nivell de risc depèn també del context en el qual el sistema es desenvolupa.
- El discurs que traspuja la PRIA està molt centrat en l'exhaustivitat de la correcció d'errors tècnics per abordar la discriminació, quan l'equitat en els sistemes d'IA depèn també d'altres factors. Les exigències de qualitat de les dades que es pretenen instaurar són ambiciosos però s'ha criticat que són impossibles de complir (p. e. bases de dades lliures d'errors i completes). Si bé és cert que la supervisió humana s'estableix com un mecanisme repetitiu, continu, que es durà a terme durant tot el cicle de vida del sistema, generarà registres i es durà a terme

de manera efectiva, per l'experiència prèvia segueix generant suspicàcies en relació amb la possibilitat que es creï un fals sentit de seguretat sense proveir realment de protecció legal.

1.7. La responsabilitat de l'acció discriminatòria no és una qüestió prou ben resolta per la normativa atès que presenta molta complexitat. Quan qui proveeix i qui usa el sistema algorítmic són la mateixa empresa o organització, és evident que la responsabilitat d'una acció discriminatòria ha de ser seva. Quan qui proveeix el sistema i qui l'usa són diferents, o bé es pot demandar per l'acció discriminatòria a qui l'usa, atès que serà qui, per imprudència, ha utilitzat el sistema algorítmic com a eina que ha tingut resultats discriminatoris, o bé es pot demandar per l'acció discriminatòria a qui el proveeix, atès que no l'ha creat amb suficient cura envers els possibles resultats discriminatoris que podia crear, o bé es pot demandar tant a qui el proveeix com a qui l'usa a la vegada.

Per una banda, la política de responsabilitat proactiva de l'RGPD pot aclarir qüestions sobre la responsabilitat dels efectes discriminatoris del sistema algorítmic, que seguint l'exemple del tractament de dades il·lícit, ha de recaure sobre l'empresa o organització que el posa en pràctica, llevat que la persona responsable de tractament l'hagi causat per manca de diligència. Tanmateix, la responsabilitat tal com està plantejada en l'RGPD es basa en el tractament de dades, que és una acció que duu a terme d'inici a fi la mateixa empresa o organització i no compta amb cap eina de suport prèviament configurada -com pot ser un sistema algorítmic.

Per l'altra banda, la PRIA clarament apunta com a responsable de les conseqüències derivades de la comercialització del sistema algorítmic la figura del proveïdor i en canvi quasi no atribueix responsabilitats a qui l'usa. La responsabilitat de la persona que l'usa vindrà derivada del coneixement o la manca de diligència en escollir una eina de presa de decisió no discriminatòria. En aquest cas, pot ser molt difícil provar que l'empresa o organització que ha emprat un sistema algorítmic discriminatori de tercers era conscient que els resultats que obtindria serien discriminatoris, i per tant, que es pot incloure en la responsabilitat de l'acció discriminatòria.

Davant d'aquestes disjuntives, la responsabilitat tal com està plantejada en la Llei 15/2022 pot aportar aclariments: l'article 27 estableix que la persona física o jurídica que

causi discriminació repararà el dany proporcionant una indemnització (...) i seran igualment responsables del dany causat les persones que contracten o prestadores de serveis quan la discriminació es produeixi en el seu àmbit d'organització i direcció i no hagin aplicat mètodes i instruments per la detecció, aplicació de mesures preventives i articulació de mesures adequades pel cessament de situacions discriminatòries (en relació amb l'article 25 de la mateixa norma). També resulta d'interès l'article 44 de la Llei 19/2020, del 30 de desembre, d'igualtat de tracte i no-discriminació que només és d'aplicació a Catalunya, que estableix que són responsables de les infraccions tipificades en la llei les persones físiques o jurídiques, públiques o privades, que, per acció o omissió, incorren en els supòsits d'infracció, i si una infracció és imputable a diverses persones i no és possible determinar el grau de participació de cadascuna en la comissió de la infracció (per exemple quin grau de participació en el resultat discriminatori té el proveïdor i quin grau de participació té qui usa el sistema), n'han de respondre solidàriament. Aquesta última apreciació és interessant en el cas que no es pogués determinar si la responsabilitat és del proveïdor o de l'empresa o organització que utilitza el sistema algorítmic, ja que en no poder-se determinar el grau de participació de cadascuna, haurien de respondre solidàriament.

1.8. Algunes garanties de la normativa europea antidiscriminació, de protecció de dades i de la PRIA, podrien ser vàlides per protegir les persones de la discriminació algorítmica (p. e. igualtat des del disseny, avaluacions d'impacte algorítmiques, documentació tècnica en virtut de la PRIA o la possibilitat de sol·licitar al tribunal que l'altra part en un litigi reveli informació que té en la seva possessió), donades certes condicions. Tanmateix, no considerem que constitueixin règims protectors de suficient entitat com per donar eines adequades i suficients a les persones que es vegin discriminades per desafiar les decisions dels sistemes algorítmics amb les quals s'ha causat la discriminació i accedir a la justícia.

2. Propostes concretes (en virtut de les conclusions generals)

2.1. Pel control de la discriminació en els sistemes algorítmics:

Per establir un control efectiu de la discriminació que poden generar els sistemes algorítmics de presa de decisió aplicats en contextos socials s'ha d'abandonar la idea que existeixen sistemes algorítmics (o persones) que no discriminin i treballar per limitar normativament quanta discriminació algorítmica és tolerable. És rellevant la concreció d'una normativa que esperi dels sistemes algorítmics l'assoliment d'uns estàndards concrets en relació amb els resultats discriminatoris.

Una aposta coherent i factible dins els límits de la regulació antidiscriminació actual (que no té com a objectiu ser una eina transformadora de les estructures socials d'opressió/subordinació, sinó que dins de les estructures ja creades, té vocació de corregir evidències flagrants puntuals de desavantatge per certs grups de població), passa per assegurar que la precisió del sistema algorítmic per tots els grups de població concrets (incloent categories de dades combinades) no superi certa taxa d'error. Concretament, establir taxes d'error tolerables i a partir d'aquestes establir controls *ex ante* i *ex post*. Per exemple, la normativa pot establir que no es poden emprar sistemes en els quals les taxes d'error per tots els grups -tenint en compte també la interseccionalitat (p. e. un grup de població que siguin dones, afroamericanes de més de 60 anys)- superi l'1%. En el mateix sentit, però establint un altre sistema de cribratge, la normativa pot exigir que el sistema algorítmic sigui un 95% precís per tots els grups de població (incloent categories protegides combinades). Quant al control *ex ante* i *ex post* (que hauria de ser fet per terceres parts independents, no per la mateixa empresa amb intenció de comercialitzar-lo o que l'està comercialitzant), si el sistema algorítmic, o bé supera les taxes d'error, o bé és tan inescrutable que no pot demostrar-se que no les supera, s'hauria d'impossibilitar la seva comercialització o imposar la seva retirada del mercat.

Tant pel que fa al control *ex ante*, com al control *ex post*, el que s'hauria de comprovar rellevant per la discriminació algorítmica és que el sistema algorítmic emprat assoleix aquests estàndards. De manera que els dos tipus de control haurien de pivotar sobre aquesta norma. El control *ex post*, en tot cas, hauria de ser periòdic i hauria de comprovar si amb el pas del temps l'algoritme, tot i canviar, segueix essent tant precís com exigeix la normativa. Si en seu judicial es demostrés que el sistema algorítmic, o bé té una taxa d'error més alta que aquella permesa, o bé és tan inescrutable que no pot demostrar-se quines són les seves taxes d'error, s'hauria de considerar discriminatori.

Havent establert que tant les persones com els sistemes algorítmics no poden no

discriminar al cent per cent, el que proposem aquí és comprendre que l'abast de la normativa és limitat, i que per poder, si més no, plantejar una regulació que pugui ser efectiva, s'ha de ser conscient de la limitació i no crear normatives atractives però finalment del tot ineficients. L'objectiu és que el sistema algorítmic ha d'assolir certs estàndards de precisió ja que no pot ser no discriminatori, de la mateixa manera que actualment a les persones no se'ls hi demana ser no discriminatòries, sinó que en la seva interacció amb tercers i en la presa de decisions que els afecten, respectin certs límits establerts per la normativa.

Problema: Tanmateix, si aquesta proposta normativa no es fa a nivell europeu, no sembla que pugui ser viable. Tant en l'RGPD com en la futura regulació de la intel·ligència artificial es pot advertir, més enllà de la finalitat protectora, l'objectiu de la UE per crear unes bases que facilitin un mercat digital únic i evitar que es continuïn produint obstacles per al mercat intern. La introducció d'una normativa en aquest sentit per part d'un estat membre pot no estar permesa atès que suposaria un obstacle a la circulació dels productes en el mercat interior de la UE. Aquesta conclusió ens ha orientat a investigar sobre les possibles eines reactives que poden proporcionar els estats membres de la UE dins les seves competències per desafiar les decisions que produeixin discriminació algorítmica.

2.2. Propostes per facilitar el desafiament de decisions que hagin produït discriminació algorítmica:

2.2.1. Creació d'una base de dades de sistemes de presa de decisió mitjançant sistemes algorítmics a escala de la UE (tal com la base de dades que es planteja en l'article 60 de la PRIA en relació amb els sistemes de risc alt descrits en l'annex III de la norma, però amb tots els sistemes de presa de decisió amb els quals s'empren algoritmes). Aquesta llista hauria de ser accessible al públic. Prèviament s'hauria d'obligar els proveïdors de sistemes d'IA que denominin el seu sistema algorítmic, de manera que es pugui reconèixer per part de les persones sobre les quals ha pres decisions. Establerta una base de dades, s'ha d'obligar les empreses, organitzacions i administracions públiques a publicar un catàleg dels sistemes algorítmics que estan utilitzant i una explicació de per a què els estan utilitzant (p. e. en una pàgina web accessible al públic general que

s'especifiqui que s'està emprant el sistema algorítmic "x" per gestionar els processos selectius de l'empresa).

2.2.2. Establiment d'una norma que faciliti l'accés a informació en poder de l'altra part en el context d'un litigi per discriminació. Ja hi ha normes en aquest sentit, com l'article 18 de la Directiva 2020/1828 (que pot ser aplicable en casos de discriminació algorítmica quan es doni col·lateralment a la il·licitud d'un tractament de dades en l'àmbit del consum), on s'estableix que els estats membres vetllaran per tal que quan una entitat habilitada hagi aportat proves raonablement disponibles, suficients per donar suport a una acció de representació, i hagi indicat altres proves que estiguin en poder de la part demandada o d'una tercera part, si així ho demana l'entitat habilitada, l'òrgan jurisdiccional o l'autoritat administrativa pugui ordenar a la part demandada o tercera part que exhibeixi les proves sol·licitades. Es va redactar, en el mateix sentit, l'article 4.b de la proposta de Directiva de la Comissió de 1988, relativa a la càrrega de la prova en l'àmbit de la igualtat de retribució, on s'establí que les parts en litigi podien disposar de qualsevol informació necessària per fer valer els seus drets en poder de l'altra part (aquest article que no va arribar al text final).

Un article en aquest sentit en la normativa antidiscriminatòria facilitaria a la part demandant obtenir informació sobre la precisió del sistema algorítmic en certs grups determinats de població o qualsevol informació pertinent per provar que s'ha produït la discriminació algorítmica. Actualment, atesa la manca d'un article en aquest sentit en l'àmbit antidiscriminació, interessa la potestat dels organismes d'igualtat nacionals per obtenir de les administracions públiques i empreses la informació que sol·licitin en el marc d'una investigació sobre un cas de discriminació.

2.2.3. Instaurem un sistema de protecció col·lectiva-institucional que instrumentalitzi els punts forts que ofereixen les normatives que regulen o poden regular col·lateralment la discriminació algorítmica per millorar les possibilitats de desafiar les decisions preses mitjançant sistemes algorítmics que puguin causar discriminació. A nivell individual, la protecció que es pugui proporcionar sempre serà insuficient.

2.2.3.1. Promoció de l'acció col·lectiva en matèria de discriminació algorítmica

L'acció col·lectiva (que podria dur-se a terme per protegir interessos col·lectius, difusos, (pluri)individuals homogenis o l'interès general i podria tenir l'objectiu tant de cessar l'acció discriminatòria com de reparar els perjudicis causats) és adequada per adreçar la discriminació algorítmica per diverses raons:

- Posa en rellevància el component grupal de la discriminació, és a dir, l'objectiu a protegir a través de l'acció col·lectiva és el grup de persones que es veu tractat més desfavorablement, permetent a través de l'acció, aconseguir que s'assoleixi el cessament de l'ús del sistema algorítmic que produeix l'acció discriminatòria (que perpetua el tracte perjudicial a un grup determinat de població).
- Millora les possibilitats d'establir un cas *prima facie*, en haver més persones legitimades per desafiar la discriminació algorítmica. També, si múltiples persones que pertanyen a un col·lectiu de població determinat que comparteix una (o diverses) categoria(es) protegida(es) plantegen que s'han vist perjudicades per un mateix sistema algorítmic, serà més fàcil evidenciar que el tracte rebut per part del sistema algorítmic és sistemàticament pitjor a causa de la pertinença a la/les categoria/es protegida/es.
- Millora les possibilitats d'establir *tertium comparationis*, atès que hi ha entitats legitimades que tenen més possibilitats d'accedir a informació sobre el sistema algorítmic que no pas les persones individualment considerades (els organismes d'igualtat nacionals).
- Atesa la dispersió i magnificació dels efectes de la discriminació algorítmica, la protecció dels interessos difusos poden tenir especial rellevància en el cas de la discriminació algorítmica. També pot tenir especial rellevància l'acció de cessació, atès que moltes persones no seran conscients de la discriminació -i per tant no demanaran reparació-, i l'objectiu principal (de les associacions i organitzacions que la promoguin) serà aconseguir que l'algoritme discriminatori no pugui seguir operant i es declari discriminatori.

A la UE no hi ha un mecanisme harmonitzat d'acció col·lectiva antidiscriminació. Únicament les directives antidiscriminació obliguen els estats membres a que associacions, organitzacions i entitats jurídiques puguin iniciar en nom de la persona demandant o en el seu suport i autorització procediments judicials i administratius. El TJUE, no obstant això, ha admès que no sigui necessària una víctima de discriminació identificada per portar un cas davant del tribunal (p. e. cas Feryn); o la suficiència d'un dany hipotètic (p. e. cas *Accept i Rete Lenford*).

Problema: En relació amb la suficiència d'un dany hipotètic, des de la nostra perspectiva cal tenir en compte que en els casos esmentats es tractava de casos molt concrets en els quals hi havia una declaració pública que evidenciava com un grup determinat de població que compartia un atribut protegit s'havia vist discriminat. En el cas de la discriminació algorítmica, no hi hauria unes declaracions que evidenciessin la discriminació sinó que directament hi hauria exclusió, i llavors, el raonament jurídic del TJUE podria no ser aplicable.

Els estats membres poden oferir més possibilitats que les establertes en les directives antidiscriminació, com és el cas d'Espanya, on sí que és possible la defensa d'interessos difusos de persones que s'han vist discriminades (art. 11 bis.2 de la LEC), ocasionant que no sigui únicament la persona discriminada que hagi de vetllar per la seva protecció, sinó que entitats amb més accés a informació i recursos, per exemple els organismes d'igualtat nacionals, puguin desafiar decisions que causin discriminació algorítmica i plantejar la cessació de l'algoritme que causa la discriminació.

La normativa i jurisprudència en matèria de protecció de dades també plantegen la possibilitat de la defensa de drets mitjançant l'acció col·lectiva. L'article 80.2 de l'RGPD possibilita que entitats, associacions i organitzacions puguin dur a terme accions de cessació en l'àmbit de la protecció de dades, sobretot en els casos on, havent-se habilitat a nivell nacional, no sigui necessari un mandat ni la vulneració de drets concrets de la persona interessada (d'acord amb la sentència del TJUE de 28 d'abril de 2022, *Meta Platforms Ireland Limited*¹⁹⁷⁴). Per tant, quan un tractament de dades il·lícit provoqui discriminació algorítmica podria ser desafiat a través d'una acció col·lectiva en aquest

¹⁹⁷⁴ TJUE (Sala Tercera), sentència de 28 d'abril de 2022, C-319/20, *Meta Platforms Ireland Limited*, anteriorment *Facebook Ireland Limited*, contra *Bundesverband der Verbraucherzentralen und Verbraucherverbände - Verbraucherzentrale Bundesverband e.V.*

sentit. No obstant això, a Espanya no s'ha aprofundit en l'acció col·lectiva de protecció de dades i només està habilitada en el cas que es doni dins d'una relació de consum. En aquests casos, es podria fer ús de l'article 18 de la Directiva 2020/1828 (transposat degudament en cada ordenament nacional) i demanar al tribunal que la part demandada exhibeixi proves sol·licitades que puguin ajudar a provar la discriminació.

Primer problema: Que l'àmbit d'incidència del tractament massiu de dades no tingui límits territorials i en canvi els requisits per la protecció dels drets en sentit col·lectiu estiguin delimitats territorialment. Que els sistemes algorítmics operin sense límits territorials afecta directament a l'efectivitat de les normatives i proteccions dels drets fonamentals.

Segon problema: L'acció col·lectiva és un instrument que plantejat per si mateix presenta certes limitacions. Les persones tenen molt poc accés a informació de rellevància, estan molt disseminades i isolades, i són poc conscients o estan poc preocupades pels riscos que planteja la intel·ligència artificial. És necessari, llavors, incentivar la institucionalització de l'acció col·lectiva, atès que no es podrà brindar una protecció efectiva sense un organisme públic que pugui, per exemple, accedir a la informació sobre l'impacte precís del sistema algorítmic sobre grups determinats de persones; centralitzar (telemàticament) les demandes de les persones afectades per un sistema algorítmic presumptivament discriminatori i gestionar i promoure una acció de cessació de l'acció discriminatòria (a través de la combinació entre la informació individual i la informació sobre la precisió del sistema algorítmic en determinats grups de població); i posar en contacte comunitats afectades i proveïdors d'aquests sistemes per col·laborar en crear sistemes més sensibles a les conseqüències discriminatòries.

2.2.3.2. Protecció col·lectiva-institucional envers la discriminació algorítmica

En l'estat actual de la qüestió, la figura dels organismes d'igualtat nacionals s'ha d'explotar ja que sota la nostra perspectiva són els millor posicionats per adoptar un rol proactiu en defensa de les persones que es vegin perjudicades mitjançant sistemes algorítmics; la PRIA no dona drets col·lectius ni estableix mecanismes d'acció col·lectiva envers sistemes algorítmics discriminatoris, però facilita l'accés a la documentació tècnica dels sistemes d'IA de risc alt per part dels organismes d'igualtat (article 64.3 i

considerant 79 en relació amb l'article 11 de la norma), possibilitant que a través d'aquesta informació es pugui evidenciar que la precisió del sistema algorítmic és comparativament millor o pitjor segons el grup determinat de població; la normativa antidiscriminació i l'RGPD possibiliten l'acció col·lectiva en matèria de discriminació algorítmica però no faciliten l'obtenció de mitjans de prova que proveeixin dels elements de comparació necessaris per establir casos de discriminació.

En conseqüència, el sistema d'obtenció de prova i tutela de les persones envers la discriminació algorítmica que veiem més factible assumeix que les persones -davant la intuïció d'haver patit discriminació per part d'un sistema algorítmic, ja sigui a través dels resultats que han obtingut per part del sistema algorítmic (seria interessant, com hem esmentat, que els sistemes algorítmics tinguessin una denominació que els identifiqués i fos públic per a quina presa de decisió s'estan emprant), ja sigui a través de la informació obtinguda exercint els drets individuals que els confereix l'RGPD-:

1. Poguessin fer un queixa a un organisme d'igualtat nacional, en el cas espanyol a l'Autoritat Independent. També es podria donar el cas que l'organisme d'igualtat tingués notícia que s'està produint discriminació algorítmica per qualsevol mitjà (arran de la informació obtinguda per part de l'organisme pròpiament o de la informació que pot obtenir a través de la documentació tècnica de sistemes d'IA de risc alt, o de la coordinació amb altres organismes) i iniciés una investigació.
2. L'organisme d'igualtat nacional obtingués, en el cas que el sistema que presumptivament està produint discriminació fos classificat de risc alt pel Reglament de la IA, la documentació tècnica presentada per la persona proveïdora del sistema, o emprés les prerrogatives d'investigació de les quals disposa per obtenir informació que esclarís si s'ha produït discriminació algorítmica.
3. A partir d'aquesta informació tècnica i amb la informació proporcionada a través de les queixes o a través de les seves facultats d'investigació, l'organisme nacional d'igualtat o bé actués com a òrgan quasi-jurisdiccional i prengués una decisió sobre el sistema algorítmic discriminatori i sancionés -si fos pertinent- l'organització o administració pública que estigués emprant el sistema algorítmic discriminatori; o bé recollís indicis clars que les persones amb una (o diverses)

categoria(es) protegida(es) per la legislació han patit discriminació *prima facie* causada per l'ús del sistema algorítmic i,

4. O bé iniciés d'ofici accions en defensa de l'interès general o d'interessos difusos per aconseguir el cessament de l'ús del sistema algorítmic que està causant discriminació (podent emprar estadístiques confeccionades pel propi organisme com element comparador o d'altres a les quals tingués accés); o bé iniciés accions de cessació o participés com a part en accions en defensa d'interessos col·lectius -si les persones perjudicades estiguessin degudament identificades-; o bé proporcionés informació a les persones perjudicades per tal que poguessin dur a terme accions en defensa dels seus interessos o d'interessos (pluri)individuals homogenis, habilitant les persones individuals implicades per sol·licitar també reparació pel dany moral causat.

Aquest rol proactiu per part de l'organisme nacional d'igualtat per proporcionar una protecció col·lectiva-institucional amb efectivitat, s'hauria de complementar amb l'administració d'una plataforma pública que facilités la gestió de queixes per discriminació algorítmica, interrelacionés persones perjudicades i pogués detectar problemàtiques sistèmiques per iniciar litigació estratègica i iniciés accions de cessació amb l'objectiu de fer cessar l'ús de sistemes algorítmics discriminatoris. També seria interessant que s'estudiessin els casos per abordar de manera unificada la discriminació interseccional, es participés en els espais de prova dels sistemes algorítmics que preveu la PRIA (art. 53), es mantingués una llista de sistemes algorítmics que s'hagin provat discriminatoris, i s'establís una coordinació estable amb les agències de supervisió de la intel·ligència artificial i de protecció de dades.

La instauració d'un sistema de protecció col·lectiva-institucional envers la discriminació algorítmica podria millorar la situació de la persona (o persones) perjudicada(es) principalment perquè donaria més oportunitats de crear un grup perjudicat que pogués desafiar la decisió discriminatòria (reprement el punt 1.1 de les conclusions), incentivaria la possibilitat de crear un grup perjudicat i atorgaria la responsabilitat de combatre la discriminació algorítmica a un organisme amb més recursos (p. e. possibilitat d'accedir a informació i proves de discriminació, o a estadístiques que poden funcionar com a element comparador), que al seu torn també podria actuar d'ofici, i tenir com a objectiu la protecció del col·lectiu perjudicat, de manera que l'abast de les seves accions podria

ser molt més ampli.

2.2.3.3. L'Autoritat Independent per a la igualtat de tracte i no discriminació -en tant que organisme d'igualtat- disposa de l'especialització i les funcions necessàries per posicionar-se com a òrgan de referència en la lluita envers la discriminació algorítmica, però ha de coordinar-se i comptar amb la cooperació tant de l'AEPD com de l'AESIA.

L'Autoritat Independent tindrà la capacitat per elaborar informes i estadístiques de caràcter periòdic, per iniciar d'ofici o a instància de tercers investigacions sobre l'existència de possibles situacions de discriminació que plantegin especial gravetat o rellevància, per exigir col·laboració necessària a administracions públiques i particulars, possibilitant l'accés a informació i dades que l'Autoritat Independent necessiti en les seves investigacions, per exercitar accions judicials en defensa dels drets derivats de la igualtat de tracte i no discriminació i per rebre queixes individuals de les persones, possibilitant l'anàlisi i identificació de similituds entre elles, establint la concurrència d'un problema sistèmic. També podrà accedir a la documentació tècnica dels sistemes de risc alt que els proveïdors estaran obligats a fer en virtut de la PRIA, estar implicada en els espais de prova dels sistemes algorítmics que planteja la PRIA amb l'objectiu de captar possibles efectes discriminatoris.

Tanmateix considerem que actualment, malgrat que es configura com una autoritat especialitzada en discriminació, no ho és en el camp de la intel·ligència artificial -com l'AESIA- ni compta amb l'experiència d'integrar la perspectiva de la protecció de drets fonamentals en un camp tècnic -com l'AEPD-, de manera que per tal d'incrementar l'efectivitat de la seva actuació en l'àmbit de la discriminació algorítmica haurà d'actuar en un marc de complementaritat amb aquestes dues agències. També dins un marc de coordinació, sobretot davant la possibilitat per part de l'AESIA i l'AEPD de detectar possibles vulneracions del dret a la no discriminació i incentivant l'intercanvi d'informació entre organismes. En aquest sentit, potser caldrà també establir marcs de cooperació per tal de disposar de tota l'especialització i les eines possibles per fer front a la discriminació algorítmica. D'aquesta manera, per exemple, si presumptivament ha ocorregut una discriminació, la cooperació entre l'Autoritat Independent i l'AESIA, pot facilitar l'accés a informació relativa a la precisió del sistema d'IA sobre determinats

grups de població per tal d'interrelacionar persones que s'han vist perjudicades i iniciar accions de cessació de l'ús del sistema d'IA en qüestió.

El nostre estudi s'ha centrat en les eines i recursos que dona actualment la normativa per fer front a la discriminació algorítmica i, davant la seva ineficiència, hem proposat una eina de protecció col·lectiva-institucional centrada sobretot en la creació d'un grup perjudicat que pugui desafiar la decisió discriminatòria produïda mitjançant sistemes algorítmics. Tanmateix aquest és un àmbit en constant evolució i queden obertes moltes qüestions que, en part, s'aniran resolent a mida que es perfili un regulació més concreta dels sistemes d'intel·ligència artificial. Tanmateix, sobre la qüestió de la discriminació algorítmica tenen una incidència directa quines exigències imperatives de transparència sobre els sistemes algorítmics s'acabaran imposant i com de difícil serà per a les persones accedir a informació rellevant sobre les decisions que s'han pres sobre elles. Addicionalment també ens preguntem fins a quin punt s'afavorirà a través de les institucions públiques la creació d'espais i mitjans de protecció per tal que les persones perjudicades s'uneixin i puguin desafiar i qüestionar les decisions a les quals s'han vist sotmeses. Finalment, està per veure si, o bé una única autoritat assumirà tant la tasca de supervisar els sistemes que emprin la IA com la protecció dels drets de les persones que es puguin veure vulnerats, o bé es crearan autoritats especialitzades segons el sector en el qual s'apliqui la IA, o bé convergiran diferents autoritats que s'aniran redefinint i nodrint de competències en matèria d'IA i anàlisi massiva de dades -aquesta seria la nostra perspectiva-, que per brindar una protecció efectiva, indefectiblement s'hauran de coordinar i cooperar entre elles.

BIBLIOGRAFIA

ACINAS GARCÍA, N., “Seguridad de los tratamientos: los análisis de riesgos, la seudonimización, el cifrado y la adhesión a códigos de conducta (Comentario al artículo 32 RGPD)”, a TRONCOSO REIGADA, A., (dir.) i VILLALBA CANO, L., (coord.), *Comentario al Reglamento General de Protección de Datos y a la Ley Orgánica de Protección de Datos Personales y Garantía de los Derechos Digitales*, Tomo I, Pamplona, Aranzadi, 2021, pàg. 2033-2042.

ADAM, A., *Artificial knowing: Gender and the thinking Machine*, Florence, KY, USA, Routledge, 1998.

ADAMS-PRASSL, J., “Regulating algorithms at work: Lessons for a ‘European approach to artificial intelligence’”, *European labour law journal*, vol.13, núm. 1, 2022, pàg. 30-50.

ADAMS-PRASSL, J., BINNS, R. i KELLY-LYTH, A., “Directly Discriminatory Algorithms”, *Modern Law Review*, vol. 86, núm. 1, 2023, pàg. 144-175.

AEPD, *Adecuación al RGPD de tratamientos que incorporan Inteligencia Artificial. Una introducción*, Madrid, 2020.

- *Naturaleza del dato de fumador como dato de salud*, informe jurídico 0129/2005, Madrid, 2005.
- *Datos de salud de personas fallecidas*, informe jurídico 0149/2008, Madrid, 2008
- *Guía práctica de análisis de riesgos en los tratamientos de datos personales sujetos al RGPD*, Madrid, 2018.
- *Orientaciones para la realización de una evaluación de impacto para la protección de datos en el desarrollo normativo*, Madrid, 18 de abril de 2023.

AGENCIA EUROPEA DE DERECHOS FUNDAMENTALES (FRA) y CONSEJO DE EUROPA, *Manual de legislación Europea contra la discriminación*, Luxemburgo, 2019.

AGUIRREZABAL GRÜNSTEIN, M. “Algunas precisiones en torno a los intereses supraindividuales (colectivos y difusos)”, *Revista Chilena de Derecho*, vol. 33, núm. 1, 2006, pàg. 69-91.

AGULLÓ AGULLÓ, D., “Los contratos de financiación de litigios por terceros (third-party funding) en España”, *Revista de Derecho Civil*, vol. IX, núm. 1, 2022, pàg. 183-231.

AJUNWA, I., “Discrimination by platforms”, *Berkeley Journal of Employment and labour law*, vol. 40, núm. 1, 2019, pàg. 1-21.

- “The paradox of automation as anti-bias intervention”, *Cardozo Law Review*, vol. 40, 2020, pàg. 1671-1742.

ALLEN, R. i MASTERS, D., “Artificial Intelligence: the right to protection from discrimination caused by algorithms, machine learning and automated decision-making”, *ERA Forum*, vol 20, núm. 4, 2020, pàg. 585-598.

- *Regulating for an equal AI: A new role for equality bodies. Meeting the new challenges to equality and non-discrimination from increased digitisation and the use of Artificial Intelligence*, European Network of Equality Bodies (EQUINET), Brussels, 2020.

ALLHUTTER, D., CECH, F., FISCHER, F., GRILL, G. i MAGER, A., “Algorithmic Profiling of Job Seekers in Austria: How Austerity Politics Are Made Effective”, *Frontiers Big Data*, vol. 3, art. 5, 2020, pàg. 1-17.

ALMADA, M., “Human intervention in automated decision-making. Toward the construction of contestable Systems”, *ICAAIL '19: Proceedings of the Seventeenth International Conference on Artificial Intelligence and Law*, 2019, pàg. 2-11.

ALPAYDIN, E., *Machine learning: the new AI*, Massachusetts, The MIT Press, 2016.

ALSTON, P., *Digital welfare states and human rights*, UN Special Rapporteur on extreme poverty and human rights, report A/74/493, New York, 2019.

ÁLVAREZ DEL CUVILLO, A., “El problema de la discriminación inversa: ¿Es posible discriminar a quienes pertenecen a los grupos sociales dominantes?”, *Trabajo, Persona*,

Derecho, Mercado, vol. 5, 2022, pàg. 187-209.

ALVAREZ VIGARAY, R., “La responsabilidad por daño moral”, *Anuario de derecho civil*, vol. 19, núm. 1, 1966, pàg. 81-116.

AMRAM, D., “The Role of the GDPR in Designing the European Strategy on Artificial Intelligence: Law-Making Potentialities of a Recurrent Synecdoche”, *Opinio Juris in Comparatione. Studies in comparative and National Law*, vol. 1, núm. 1, 2020, pàg. 73-95.

ANANNY, M., i CRAWFORD, K., “Seeing without Knowing: Limitations of the Transparency Ideal and Its Application to Algorithmic Accountability”, *New Media and Society*, vol. 20, núm. 3, 2018, pàg. 973-989.

ANDRUS, M. *et al.*, “What We Can’t Measure, We Can’t Understand: Challenges to Demographic Data Procurement in the Pursuit of Fairness”, *Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency*, 2021, pàg. 249-260.

AÑÓN ROIG, M. J., “Principio antidiscriminatorio y determinación de la desventaja”, *Isonomía: Revista de teoría y filosofía del derecho*, núm. 39, 2013, pàg. 127-157.

- “Desigualdades algorítmicas: conductas de alto riesgo para los derechos humanos”, *Derechos y libertades*, núm 47, època II, 2022, pàg. 17-49.

ARABADJIEVA, K. i ZWYSEN, W., *Gender inequality in performance-related pay: a gap in the EU equal pay agenda*, ETUI Research Paper - Policy Brief, Brussels, February 2022.

ARCOS VARGAS, M., “La no discriminación en el derecho derivado de la Unión Europea”, a MORALES ORTEGA, J. M., (coord.), *Realidad social y discriminación: Estudios sobre diversidad e inclusión laboral*, 2022, pàg. 17-41.

ARENAS RAMIRO, M., “El derecho de acceso y las condiciones generales de ejercicio de los derechos (Comentario al artículo 15 RGPD y a los artículos 12 y 13 LOPDGDD)”, a TRONCOSO REIGADA, A., (dir.) i VILLALBA CANO, L., (coord.), *Comentario al Reglamento General de Protección de Datos y a la Ley Orgánica de Protección de Datos Personales y Garantía de los Derechos Digitales*, Tomo I, Pamplona, Aranzadi, 2021,

pàg. 1437-1490.

- “El derecho de oposición (Comentario al artículo 21 RGPD y al artículo 18 LOPDGDD)”, a TRONCOSO REIGADA, A., (dir.) i VILLALBA CANO, L., (coord.), *Comentario al Reglamento General de Protección de Datos y a la Ley Orgánica de Protección de Datos Personales y Garantía de los Derechos Digitales*, Tomo I, Pamplona, Aranzadi, 2021, pàg. 1701-1724.

ARTICLE 29 DATA PROTECTION WORKING PARTY, *Opinion 4/2007 on the concept of personal data*, 01248/07/EN, WP 136, Brussels, 20 June 2007.

- *Guidelines on Data Protection Impact Assessment (DPIA) and determining whether processing is “likely to result in a high risk” for the purposes of Regulation 2016/679*, 17/EN, WP 248 rev.01, Brussels, 4 October 2017.
- *Guidelines on Automated individual decision-making and Profiling for the purposes of Regulation 2016/679*, 17/EN, WP 251, Brussels, 6 February 2018.

ATREY, S., “Comparators in intersectional discrimination”, *Legal Studies*, vol. 38, núm. 3, 2018, pàg. 379-395.

AUTORITAT CATALANA DE PROTECCIÓ DE DADES, *Intel·ligència artificial. Decisions Automatitzades a Catalunya*, Barcelona, 2020.

- *Dictamen en relació amb la consulta formulada per un organisme en relació amb la instal·lació de càmeres de videovigilància pel control del trànsit en el casc urbà*, Barcelona, 11 de febrer de 2021.

BAEZA-YATES, R., “Bias on the Web”, *Communications of the ACM*, vol. 61, núm. 6, 2018, pàg. 54-61.

BAKER, J., DENG, L., GLASS, J., KHUDANPUR, S., LEE, C. H., MORGAN, N. i O'SHGUGHNESSY, D., “Research Developments and Directions in Speech Recognition and Understanding. Part 1”, *IEEE Signal Processing Magazine*, vol. 26, núm. 3, 2009, pàg. 75-80.

BALAYN, A. i GÜRSES, S., *Beyond debiasing. Regulating AI and Its Inequalities*, European Digital Rights (EDRi), Brussels, 2021.

BAROCAS, S. i SELBST, A. D., “Big Data’s Disparate Impact”, *California Law Review*, vol. 104, 2016, pàg. 671-732.

BARRÈRE, M.A., “Iusfeminismo y derecho antidiscriminatorio: hacia la igualdad por la discriminación”, a MESTRE i MESTRE, R.M., (coord.), *Mujeres, derechos y ciudadanías*, Valencia, Tirant Lo Blanch, 2008, pàg. 45-71.

- “La igualdad de género desde el activismo de las profesiones jurídicas”, *Revista Vasca de Administración Pública*, vol. 1, núm. especial 99-100, 2014, pàg. 513-528.
- “Filosofías del Derecho Positivo ¿Qué Derecho y qué discriminación? Una visión contrahegemónica del Derecho Antidiscriminatorio”, *Anuario de Filosofía del Derecho*, vol. XXXIV, 2018, pàg. 11-42.

BARRÈRE, M.A. i MORONDO, D., “Subordinación y discriminación interseccional: elementos para una teoría del derecho antidiscriminatorio”, *Anales de la Cátedra Francisco Suárez*, núm. 45, 2011, pàg. 15-42.

BARROS VALE, S. i ZANFIR-FORTUNA, G., *Automated Decision-Making Under the GDPR: Practical Cases from Courts and Data Protection Authorities*, Future of Privacy Forum, Washington, DC, May 2022.

BARTLETT, K., “Feminist legal methods”, *Harvard Law Review*, vol. 103, núm. 4, 1990, pàg. 829-888.

BELJIN, S., “Rights in EU Law”, a PRECHAL, S. i VAN ROERMUND, B., (eds.), *The Coherence of EU Law. The Search of Unity in Divergent Concepts*, Oxford, Oxford University Press, 2008, pàg. 91-122.

BELL, M., *Anti-Discrimination Law and the European Union. Oxford Studies in European Law*, Oxford, Oxford University Press, 2002.

- “Constitutionalization and EU Employment Law”, a MICKLITZ, H-W., (ed.), *Constitutionalization of European private law: XXII/2*, Oxford, Oxford University Press, 2014, pàg. 137-169.

BENDER, E. M., i FRIEDMAN, B., “Data Statements for Natural Language Processing:

Toward Mitigating System Bias and Enabling Better Science”, *Transactions of the Association for Computational Linguistics*, vol. 6, 2018, pàg. 587-604.

BENEDI LAHUERTA, S., “Ethnic discrimination, discrimination by association and the Roma community: CHEZ”, *Common Market Law Review*, vol. 53, núm. 3, 2016, pàg. 797-818.

- “Enforcing EU Equality Law through Collective Redress: Lagging behind?”, *Common Market Law Review*, vol. 55, núm. 3, 2018, pàg. 783-818.
- “Equality bodies: advancing towards more responsive designs?”, *International Journal of law in context*, vol. 17, núm. 3, 2021, pàg. 390-412.

BENJAMIN, R., *Race After Technology: Abolitionist Tools for the New Jim Code*, Medford, MA, Polity, 2019.

BENNET MOSES, L. i CHAN, J., “Algorithmic prediction in policing: assumptions, evaluation, and accountability”, *Policing and Society*, vol. 28, núm. 7, 2018, pàg. 806-822.

BERENDT, B., “The AI Act Proposal: Towards the next transparency fallacy? Why AI regulation should be based on principles based on how algorithmic discrimination Works”, a BMJV i ROSTALSKI, F., (eds.), *Künstliche Intelligenz - Wie gelingt eine vertrauenswürdige Verwendung in Deutschland und Europa?*, Tübingen, Mohr Siebeck, 2022, pàg. 31-52.

BERTAIL, P., BOUNIE, D., CLÉMENÇON, S. i WAELBROECK, P., “Algorithmes: biais, discrimination et équité”, Télécom ParisTech, 2019.

BIETTI, E., “From Ethics Washing to Ethics Bashing: A Moral Philosophy View on Tech Ethics”, *Journal of Social Computing*, vol. 2, núm. 3, 2021, pàg. 266-283.

BIGMAN, Y. E., WILSON, D., ARNESTAD, M. N. WAYTZ, A. i GRAY, K., “Algorithmic Discrimination Causes Less Moral Outrage Than Human Discrimination”, *Journal of Experimental Psychology: General*, vol. 152, núm. 1, 2023, pàg. 4-27.

BILBAO, J.M., “Prohibición de discriminación y las relaciones entre particulares”, *Teoría y realidad constitucional*, núm. 18, 2006, pàg. 147-189.

BINNS, R., “Data Protection Impact Assessments: A Meta-Regulatory Approach”, *International Data Privacy Law*, vol. 7, núm. 1, 2017, pàg. 22-35.

- “Fairness in machine learning: Lessons from political philosophy”, *Proceedings of the 1st Conference on Fairness, Accountability and Transparency*, PMLR, vol. 81, 2018, pàg. 149-159.
- “On the Apparent Conflict Between Individual and Group Fairness”, *FAT* '20: Proceedings of the 2020 Conference on Fairness, Accountability, and Transparency*, 2020, pàg. 514-524.

BIRD, S. KENTHAPADI, K., KICIMAN, E. i MITCHELL, M., “Fairness-Aware Machine Learning: Practical Challenges and Lessons Learned”, *Proceedings of the Twelfth ACM International Conference on web search and data mining*, 2019, pàg. 834-835.

BLUME, P., “The myths pertaining to the proposed General Data Protection Regulation”, *International Data Privacy Law*, vol. 4, núm. 4, 2014, pàg. 269-273.

BOIX PALOP, A., “Los algoritmos son reglamentos: la necesidad de extender las garantías propias de las normas reglamentarias a los programas empleades por la administración para la adopción de decisiones”, *Revista de Derecho Público: Teoría y Método*, vol. 1, 2020, pàg. 223-270.

BOIX PALOP, A. i SORIANO ARNANZ, A. “Transparencia y control del uso de la inteligencia artificial por las administraciones públicas”, a BALAGUER CALLEJÓN, F. i COTINO HUESO, L. (coords), *Derecho Público de la Inteligencia Artificial*, Zaragoza, Fundación Manuel Giménez Abad, 2023, pàg. 247-283.

BRENNAN-MARQUEZ, K. i HENDERSON, S., “Artificial Intelligence and Role-Reversible Judgment”, *Journal of Criminal Law and Criminology*, vol. 109, núm. 2, 2019, pàg. 137-164.

BRKAN, M., “Do Algorithms Rule the World? Algorithmic Decision-Making and Data Protection in the Framework of the GDPR and Beyond”, *International Journal of Law and Information Technology*, vol. 27, núm. 2, 2019, pàg. 91–121.

BUOLAMWINI, J. i GEBRU, T., “Gender Shades: Intersectional Accuracy Disparities in Commercial Gender Classification”, *Proceedings of the 1st Conference on Fairness, Accountability, and Transparency*, vol. 81, 2018, pàg. 77–91.

BURRELL J., “How the machine ‘thinks’: Understanding opacity in machine learning algorithms”, *Big Data & Society*, vol. 3, núm. 1, 2016, pàg. 1-12.

BUTTARELLI, G., *Assessing the necessity of measures that limit the fundamental right to the protection of personal data. A Toolkit*, European Data Protection Supervisor, Brussels, 11 April 2017.

CABALLÉ-PÉREZ, M. *et al.*, “El quebrantamiento de las órdenes de protección en violencia de género: análisis de los indicadores de riesgo mediante el formulario vpr4.0”, *Anuario de psicología jurídica*, núm. 30, 2020, pàg. 63-72.

CABEZA PEREIRO, J., “La discriminación retributiva por razón de sexo como paradigma de la discriminación sistèmica”, *Lan Harremanak-Revista de relaciones laborales*, núm. 25, 2011, pàg. 79-98.

CAHN, C., “Court of Justice of the EU Rules Collective and Inaccessible Electrical Metres Discriminate against Roma: CHEZ Razpredelenie Bulgaria AD v. Komisia za zashtita ot diskriminatsia (C83/14)”, *European Journal of Migration and Law*, vol. 18, núm. 1, 2016, pàg. 112-125.

CALDERS, T. i ŽLIOBAITĚ, I., “Why Unbiased Computational Processes Can Lead to Discriminative Decision Procedures”, *Discrimination and privacy in the information society*, vol. 3, núm. 20, 2013, pàg. 43-57.

CANTARINI GUERRA, P., “Governanza algorítmica, explicación por medio del diseño y justicia del diseño”, *Anales de la Cátedra Francisco Suárez*, núm. 57, 2023, pàg. 121-141.

CAPMANY, M. A., *Cartes impertinents*, Barcelona, Horsori, 2009.

CASTANEDA, J., JOVER, A., CALVET, L., YANES, S., JUAN, A.A. i SAINZ, M., "Dealing with Gender Bias Issues in Data-Algorithmic Processes: A Social-Statistical Perspective", *Algorithms*, vol. 15, núm. 9, 2022, pàg. 1-16.

CASTAÑER CODINA, J., “La evaluación de la solvencia de las personas mediante el uso de algoritmos” a HUERGO LORA, A.J., (dir.) i DÍAZ GONZÁLEZ, G. M., (coord.), *La regulación de los algoritmos*, Pamplona, Aranzadi, 2020, pàg. 255-272.

CASTETS-RENARD, C., “Accountability of Algorithms in the GDPR and beyond: A European Legal Framework on Automated Decision-Making”, *Fordham Intellectual Property, Media and Entertainment Law Journal*, vol. 30, núm. 1, 2019, pàg. 91-137.

CAZURRO BARAHONA, V., “La supervisión de los códigos de conducta (Comentario al artículo 41 RGPD)”, a TRONCOSO REIGADA, A., (dir.) i VILLALBA CANO, L., (coord.), *Comentario al Reglamento General de Protección de Datos y a la Ley Orgánica de Protección de Datos Personales y Garantía de los Derechos Digitales*, Tomo I, Pamplona, Aranzadi, 2021, pàg. 2409-2420.

CERRILLO I MARTÍNEZ, A., “Com obrir les caixes negres de les administracions públiques? Transparència i rendició de comptes en l'ús dels algoritmes”, *Revista Catalana de Dret Públic*, núm. 58, 2019, pàg. 13-28.

- “¿Son fiables las decisiones de las Administraciones públicas adoptadas por algoritmos?”, *European Review of Digital Administration & Law - Erdal*, vol. 1, núm. 1-2, 2020, pàg. 17-36.

CHEN, M. *et al.*, “Disease prediction by Machine learning over big data from healthcare communities”, *IEEE Access*, vol. 5, 2017, pàg. 8869-8879.

CHEN, L., MA, R., HANNÁK, A. i WILSON, C., “Investigating the Impact of Gender on Rank in Resume Search Engines”, *Annual Conference of the ACM Special Interest Group on Computer Human Interaction*, 2018, pàg. 1-14.

CHOCRÓN GIRÁLDEZ, A. M., “El marco jurídico de la Directiva 2020/1828 en relación con la tutela judicial de los intereses colectivos de los consumidores”, *Cuadernos de derecho transnacional*, vol. 14, núm. 2, 2022, pàg. 275-293.

CHOPIN, I. i GERMAINE, C., *A comparative analysis of non-discrimination law in Europe 2022*, European network of legal experts in gender equality and non-discrimination, Brussels, 2022.

CHOULDECHOVA, A., “Fair prediction with disparate impact: A study of bias in recidivism prediction instruments”, *Big data*, vol. 5, núm. 2, 2017, pàg. 153–163.

CITRON, D. K., “Technological Due Process”, *Washington University Law Review*, vol. 85, núm. 6, 2008, pàg. 1249-1313.

CITRON, D. K. i PASQUALE, F., “The Scored Society: Due Process for Automated Predictions”, *Washington Law Review*, vol. 89, núm. 1, 2014, pàg. 1-33.

CLARKE, R., “Privacy Impact Assessment: Its Origins and Development”, *Computer Law and Security Review*, vol. 25, núm. 2, 2009, pàg. 123-135.

COBBE, J. “Administrative law and the machines of government: judicial review of automated public-sector decision making”, *Legal Studies*, vol. 39, núm. 4, 2019, pàg. 636-655.

COHEN J. E., “Turning Privacy Inside Out”, *Theoretical Inquiries in Law*, vol. 20, núm. 1, 2019, pàg. 1-31.

COMISIÓN EUROPEA, *Una Unión de la igualdad: Estrategia para la Igualdad de Género 2020-2025*, COM(2020) 152 final, Bruselas, 5 de marzo de 2020.

COMISIÓN EUROPEA CONTRA EL RACISMO Y LA INTOLERANCIA (ECRI), *Recomendación núm. 2 de política general de la ECRI: los órganos especializados en la lucha contra el racismo, la xenofobia, el antisemitismo y la intolerancia a nivel nacional*, Estrasburgo, 13 de junio de 1997.

- *Recomendación núm. 2 de política general de la ECRI: organismos de promoción de la igualdad para luchar contra el racismo y la intolerancia a nivel nacional*, Adoptada el 7 de diciembre de 2017, CRI(2018)06, Estrasburgo, 27 de febrero de 2018.
- *Informe de la ECRI sobre España. Quinto ciclo de supervisión*, Adoptado el 5 de diciembre de 2017, CRI(2018)2, Estrasburgo, 27 de febrero de 2018.

COMMISSION OF THE EUROPEAN COMMUNITIES, *First Report on the Implementation of the Data Protection Directive (95/46/EC)*, COM (2003). 265 final, Brussels, 15 may 2003.

COMMISSIONER FOR HUMAN RIGHTS, *Opinion of the Commissioner for Human Rights on national structures for promoting equality*, CommDH(2011)2, Council of Europe, Strasbourg, 21 March 2011.

COMMITTEE OF MINISTERS OF THE COUNCIL OF EUROPE, *Recommendation CM/Rec(2020)1 of the Committee of Ministers to Member States on the human rights impacts of algorithmic Systems*, Council of Europe, Strasbourg, 8 April 2020.

COMMITTEE ON ARTIFICIAL INTELLIGENCE (CAI), *Revised Zero Draft (Framework) convention on artificial intelligence, human rights, democracy and the rule of law*, CAI(2023)01, Council of Europe, Strasbourg, 6 January 2023.

CONSEIL DE L'EUROPE i COUR EUROPÉENNE DES DROITS DE L'HOMME, *Guide pratique sur la recevabilité*, Strasbourg, 31 août 2022.

CONSULTATIVE COMMITTEE OF THE CONVENTION FOR THE PROTECTION OF INDIVIDUALS WITH REGARD TO AUTOMATIC PROCESSING OF PERSONAL DATA, *Guidelines on Facial Recognition*, T-PD(2020)03rev4, Strasbourg, 28 January 2021.

COROMINAS BACH, S., *La legitimación activa en las acciones colectivas*, Tesi doctoral, Universitat de Girona, Girona, 2015.

COTINO HUESO, L., “Riesgos e impactos del big data, la inteligencia artificial y la robótica. enfoques, modelos y principios de la respuesta del derecho”, *Revista General de Derecho Administrativo*, núm. 50, 2019, pàg. 1-37.

- “Nuevo paradigma en la garantías de los derechos fundamentales y una nueva protección de datos frente al impacto social y colectivo de la inteligencia artificial”, a COTINO HUESO, L., (dir.) i BAUZA REILLY, M., (coord.), *Derechos y garantías ante la inteligencia artificial y las decisiones automatizadas*, Navarra, Aranzadi, 2022, pàg. 69-105.
- "Qué concreta transparencia e información de algoritmos e inteligencia artificial es la debida", *Revista Española de la Transparencia (RET)*, núm. 16, 2023, pàg. 17-63.

- “Discriminación, sesgos e igualdad de la inteligencia artificial en el sector público”, a GAMERO CASADO, E., (dir.) i PÉREZ GUERRERO, F. L., (coord.), *Inteligencia artificial y sector público. Retos, límites y medios*, Valencia, Tirant lo Blanch, 2023, pàg. 257-351.
- “Los usos de la inteligencia artificial en el sector público, su variable impacto y categorización jurídica”, 2023, *Revista Canaria de Administración Pública*, núm. 1, 2023, pàg. 211-241.

COUNCIL OF EUROPE, *The protection of individuals with regard to automatic processing of personal data in the context of profiling, Recommendation CM/Rec(2010)13 and explanatory memorandum*, Strasbourg, 23 November 2010.

COUNCIL OF EUROPE PARLIAMENTARY ASSEMBLY, *Resolution 2054 (2015) Equality and non-discrimination in the access to justice*, Strasbourg, 24 April 2015.

CRAWFORD, K., *Atlas of AI*, New Haven and London, Yale University Press, 2021.

CRAWFORD, K. i SCHULTZ, J., “Big data and due process: Toward a framework to redress predictive privacy harms”, *Boston College Law Review*, vol. 55, núm. 1, 2014, pàg. 93-128.

CROWLEY, N., *Equality bodies making a difference*, European network of legal experts in gender equality and non-discrimination, Brussels, 2018.

DATTA, A., TSCHANTZ, M. C. i DATTA, A., “Automated Experiments on Ad Privacy Settings: A Tale of Opacity, Choice and Discrimination”, *Proceedings on Privacy Enhancing Technologies*, 2015, pàg. 92-111.

DECAMP, M. i TILBURT, J. C., “Why we cannot trust artificial intelligence in medicine?” *The Lancet-Digital Health*, vol. 1, núm. 8, 2019, pàg. e-390.

DÉFENSEUR DES DROITS, *Avis du Défenseur des droits n°23-03*, Paris, 23 février 2023.

DÉFENSEUR DES DROITS (DDD) ET COMMISSION NATIONALE INFORMATIQUE ET LIBERTÉS (CNIL), *Algorithms: Preventing automation of discrimination*, Paris, 2020.

DE GREGORIO, G., *Digital Constitutionalism in Europe. Reframing Rights and Powers in the Algorithmic Society*, Cambridge, Cambridge University Press, 2022.

DELICOSTOPOULOS, J. S., “Towards European Procedural Primacy in National Legal Systems”, *European Law Journal*, vol. 9, núm. 5, 2003, pàg. 599-613.

DEL PESO NAVARRO, E. i DEL PESO RUIZ, M., “El control interno y la auditoría de los tratamientos en el marco del RGPD (Comentario al artículo 32 RGPD)”, a TRONCOSO REIGADA, A., (dir.) i VILLALBA CANO, L., (coord.), *Comentario al Reglamento General de Protección de Datos y a la Ley Orgánica de Protección de Datos Personales y Garantía de los Derechos Digitales*, Tomo I, Pamplona, Aranzadi, 2021, pàg. 2043-2067.

DESAI, D. R. i KROLL, J. A., “Trust but Verify: A Guide to Algorithms and the Law”, *Harvard Journal of Law and Technology*, vol. 31, 2017, pàg. 1-64.

DE STEFANO, V. i TAES, S., “Algorithmic management and collective bargaining”, *Transfer: European Review of Labour and Research*, vol. 29, núm. 1, 2023, pàg. 21-36.

DIAKOPOULOS, N., “Algorithmic accountability: Journalistic investigation of computational power structures”, *Digital Journalism*, vol. 3, núm. 3, 2015, pàg. 398-415.

- “Accountability in algorithmic decision making”, *Communications of the ACM*, vol. 59, núm. 2, 2016, pàg. 56-62.

DÍAZ DÍAZ, E., “El derecho de supresión. El derecho al olvido (Comentario al artículo 17 RGPD y artículo 15 LOPDGD)”, a TRONCOSO REIGADA, A., (dir.) i VILLALBA CANO, L., (coord.), *Comentario al Reglamento General de Protección de Datos y a la Ley Orgánica de Protección de Datos Personales y Garantía de los Derechos Digitales*, Tomo I, Pamplona, Aranzadi, 2021, pàg. 1561-1604.

DÍAZ GONZÁLEZ, G.M., “Algoritmos y actuación policial: la policía predictiva”, a HUERGO LORA, A.J., (dir.) i DÍAZ GONZÁLEZ, G. M., (coord.), *La regulación de los algoritmos*, Pamplona, Aranzadi, 2020, pàg. 181-234.

DIVER, L., “Digisprudence: the design of legitimate code”, *Law, Innovation and Technology*, vol. 13, núm. 2, 2021, pàg. 325–354.

DOUGAN, M., “Who Exactly Benefits from the Treaties? The Murky Interaction Between Union and National Competence over the Capacity to Enforce EU Law”, a LEINARTE, E. i ODUDU, O., (eds), *Cambridge Yearbook of European Legal Studies*, vol. 12, 2010, pàg. 73-120.

DOLEAC, J.L. i HANSEN, B., “The unintended consequences of “ban the box” statistical discrimination and employment outcomes when criminal stories are hidden”, *Journal of Labour Economics*, vol. 38, núm. 2, 2020, pàg. 321-374.

DURÁN CARDO, B., “Responsable del tratamiento (Comentario al artículo 4.7 RGPD)”, a TRONCOSO REIGADA, A., (dir.) i VILLALBA CANO, L., (coord.), *Comentario al Reglamento General de Protección de Datos y a la Ley Orgánica de Protección de Datos Personales y Garantía de los Derechos Digitales*, Tomo I, Pamplona, Aranzadi, 2021, pàg. 637-665.

- “La responsabilidad del responsable del tratamiento y las obligaciones generales del responsable. El cumplimiento proactivo de la normativa de protección de datos: de la teoría a la práctica (Comentario al artículo 24 RGPD y al artículo 28 LOPDGDD)”, a TRONCOSO REIGADA, A., (dir.) i VILLALBA CANO, L., (coord.), *Comentario al Reglamento General de Protección de Datos y a la Ley Orgánica de Protección de Datos Personales y Garantía de los Derechos Digitales*, Tomo I, Pamplona, Aranzadi, 2021, pàg. 1777-1812.

EBERS, M., “Standardizing AI. The Case of the European Commission’s Proposal for an Artificial Intelligence Act”, a DIMATTEO, L. A., CANNARSA, M. i PONCIBÒ, C., (eds.), *The Cambridge Handbook of Artificial Intelligence: Global Perspectives on Law and Ethics*, Cambridge, Cambridge University Press, 2022, pàg. 321-344.

EDELMAN, L. B., *Working Law: Courts, Corporations, and Symbolic Civil Rights*, Chicago, University of Chicago Press, 2016.

EDWARDS, L., *Regulating AI in Europe: four problems and four solutions*, Ada Lovelace Institute, London, March 2022.

EDWARDS, L. i VEALE, M., “Enslaving the Algorithm: From a “Right to an Explanation” to a “Right to Better Decisions”?”, *IEEE Security and Privacy*, vol. 16,

núm. 3, 2018, pàg. 46-54.

EGUÍLUZ CASTAÑEIRA, J.A., “Desafíos y retos que plantean las decisiones automatizadas y los perfilados para los derechos fundamentales”, *Estudios de Deusto*, vol. 68, núm. 2, 2020, pàg. 325-367.

ERICKSON, J., “Racial Impact Statements: Considering the Consequences of Racial Disproportionalities in the Criminal Justice System”, *Washington Law Review*, vol. 89, núm. 4, pàg. 1425-1465.

ESTEVE GARCÍA, F., “Las directivas europeas contra la discriminación racial y la creación de organismos especializados para promover la igualdad. Análisis comparativo de su transposición en España y en Francia”, *Revista de Derecho Constitucional Europeo*, año 5, núm. 10, 2008, 1-42.

EUBANKS, V., *Automating Inequality: How high-Tech Tools Profile, Police, and Punish the Poor*; New York, St. Martin's Press, 2018.

EUROPEAN COMMISSION, *Public Consultation: Towards a Coherent European Approach to Collective Redress*, SEC(2011)173 final, Brussels, 4 February 2011.

- *Joint Report on the application of Council Directive 2000/43/EC of 29 June 2000 implementing the principle of equal treatment between persons irrespective of racial or ethnic origin ('Racial Equality Directive') and of Council Directive 2000/78/EC of 27 November 2000 establishing a general framework for equal treatment in employment and occupation ('Employment Equality Directive')*, COM (2014) 2 final, Brussels, 17 January 2014.
- *White paper on artificial Intelligence. A European Approach to excellence and trust*, COM(2020) 65 final, Brussels, 19 February 2020.
- *Equality bodies and the implementation of the Commission Recommendation on standards for equality bodies Accompanying the document Report from the Commission to the European Parliament and Council on the application of Council Directive 2000/43/EC implementing the principle of equal treatment between persons irrespective of racial or ethnic origin ('the Racial Equality Directive') and of Council Directive 2000/78/EC establishing a general*

framework for equal treatment in employment and occupation ('the Employment Equality Directive'), COM(2021) 139 final, Brussels, 19 May 2021.

EUROPEAN DATA PROTECTION BOARD, *Guidelines 4/2019 on Article 25 Data Protection by Design and by Default*, Brussels, 20 October 2020.

EUROPEAN DATA PROTECTION BOARD i EUROPEAN DATA PROTECTION SUPERVISOR, *Joint Opinion 5/2021 on the Proposal for a Regulation of the European Parliament and of the Council Laying down Harmonised Rules on Artificial Intelligence (Artificial Intelligence Act)*, Brussels, 18 June 2021.

EUROPEAN DATA PROTECTION SUPERVISOR, *Opinion 3/2015 (with addendum), Europe's big opportunity. EDPS recommendations on the EU's options for data protection reform*, Brussels, 9 October 2015.

EUROPEAN NETWORK OF EQUALITY BODIES (EQUINET), *Strategic Litigation Handbook*, Brussels, 2017.

EUROPEAN PARLIAMENT, *European Parliament resolution of 25 March 2021 on the Commission evaluation report on the implementation of the General Data Protection Regulation two years after its application*, 2020/2717(RSP), Brussels, 25 March 2021.

EUROPEAN SUPERVISION AUTHORITIES (ESAs), *Joint Committee Final Report on Big Data*, JC/2018/04, Brussels, 15 March 2018.

EUROPEAN UNION AGENCY FOR FUNDAMENTAL RIGHTS (FRA), *Fundamental Rights Report 2018*, Luxembourg, 6 June 2018.

- *Bias in Algorithms - Artificial Intelligence and Discrimination*, Luxembourg, 8 December 2022.

FARKAS, L., "Limited enforcement possibilities under European anti-discrimination legislation -a case study of procedural novelties: actio popularis action in Hungary", *Erasmus Law Review*, vol. 3, núm. 3, 2010, pàg. 181-196.

- "Collective actions under European anti-discrimination law", *European Anti-Discrimination Law Review*, vol. 19, 2014, pàg. 25-40.

- *The Meaning of Racial or Ethnic Origin in EU Law: Between Stereotypes and Identities*, European Network of Legal Experts in Gender Equality and Non-Discrimination, Brussels, 2017.

FARKAS, L. i O'DEMPSEY, D., *How to present a discrimination claim: handbook on seeking remedies under the EU non-discrimination directives*, European Network of legal experts in the Non-discrimination field, Brussels, 2011.

FARKAS, L. i O'FARREL, O., *Reversing the burden of proof: practical dilemmas at the European and national level*, European Network of legal experts in the Non-discrimination field, Brussels, 2015.

FELDMAN, M., FRIEDLER, S., MOELLER, J., SCHEIDEGGER, C. i VENKATASUBRAMANIAN, S., "Certifying and Removing Disparate Impact", *Proceedings of the 21st ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 2015, pàg. 1-28.

FERNÁNDEZ SÁNCHEZ, S., "Frank, el algoritmo consciente de Deliveroo. Comentario a la Sentencia del Tribunal de Bolonia 2949/2020, de 31 de diciembre", *Revista de Trabajo y Seguridad Social. CEF*, núm. 457, 2021, pàg. 179-193.

FLORIDI, L., "The End of an Era: from Self-Regulation to Hard Law for the Digital Industry", *Philosophy and Technology*, vol. 34, núm. 4, 2021, pàg. 619-622.

FOUCAULT, M., *Microfísica del poder*, Edición y traducción de VARELA J. y ÁLVAREZ-URÍA, F., Madrid, Ediciones la Piqueta, 1979.

FREDMAN, S., *Discrimination Law*, Oxford, Oxford University Press, 2011.

FRIEDLER, A., SCHEIDEGGER, C. i VENKATASUBRAMANIAN, S., "The (Im)possibility of Fairness: Different Value Systems Require Different Mechanisms For Fair Decision Making", *Communications of the ACM*, vol. 64, núm. 4, 2017, pàg. 136-143.

FROOMKIN, A.M., "Regulating Mass Surveillance as Privacy Pollution: Learning from Environmental Impact Statements", *University of Illinois Law Review*, vol. 2015, núm. 5, 2015, pàg. 1713-1790.

GALLAND J. P., “The Difficulties of Regulating Markets and Risks in Europe through Notified Bodies”, *European Journal of Risk Regulation*, vol. 4, núm. 3, 2013, pàg. 365-373.

GAMPER, D., *Las mejores palabras. De la libre expresión*, Barcelona, Anagrama, 2019.

GARCÍA ÁLVAREZ, L., “La aplicación en materia ambiental de las acciones de representación para la protección de los intereses colectivos de los consumidores. A propósito de la Directiva (UE) 2020/1828”, *Revista Ítalo-Española de Derecho Procesal*, vol. 2, 2020, pàg. 58-110.

GARNICA MARTÍN, J. F., “Artículo 11. Legitimación para la defensa de derechos e intereses de consumidores y usuarios”, a FERNÁNDEZ-BALLESTEROS LÓPEZ, M.A., RIFÁ SOLER, J.M., i VALLS GOMBAU, J.F. (coords.), *Comentarios a la nueva Ley de Enjuiciamiento Civil*, Tomo I, Barcelona, Atelier, 2001, pàg. 159-173.

GARRIGA DOMÍNGUEZ, A., “La elaboración de perfiles y su impacto en los derechos fundamentales. Una primera aproximación a su regulación en el Reglamento General de Protección de Datos de la Unión Europea”, *Derechos y Libertades*, núm. 38, Época II, 2018, pàg. 107-139.

GEBRU, T., MORGENSTERN, M., VECCHIONE, B., WORTMAN VAUGHAN, J., WALLACH, H., DAUMÉ III, H. i CRAWFORD, K., “Datasheets for datasets”, *Communications of the ACM*, vol. 64, núm. 12, 2021, pàg. 86-92.

GELLERT, R. *et al.*, “A comparative analysis of anti-discrimination and data protection legislations”, a CUSTERS, B., CALDERS, T., SCHERMER, B i ZARSKY, T., (eds.), *Discrimination and privacy in the information society: data mining and profiling in large databases*, Berlin, Springer, 2013, pàg. 61-90.

GERARDS, J. i XENIDIS, R., *Algorithmic Discrimination in Europe: Challenges and Opportunities for Gender Equality and Non-Discrimination Law*, European Commission, Brussels, 2021.

GERARDS, J. i ZUIDERVEEN, F., “Protected Grounds and the System of non-discrimination law in the context of algorithmic decision-making and artificial intelligence”, *Colorado Technology Law Journal*, vol. 20, núm. 1, 2022, pàg. 3-56.

GILLIS, T.B. i SPIESS, J.L., “Big Data and Discrimination”, *University of Chicago Law Review*, vol. 86, 2019, pàg. 459-487.

GOLDBERG, S., “Discrimination by Comparison”, *Yale Law Journal*, vol. 120, 2011, pàg. 728-812.

GONZÁLEZ CAZORLA, F., “Daño moral colectivo en el derecho del consumo chileno: situación actual y proyecciones”, *Latin American Legal Studies*, vol. 8, 2021, pàg. 122-146.

GOODMAN, B. W., “A Step Towards Accountable Algorithms?: Algorithmic Discrimination and the European Union General Data Protection” *29th Conference on Neural Information Processing Systems (NIPS)*, 2016, pàg. 1-9.

GOODMAN, B., i FLAXMAN, S., “European Union Regulations on Algorithmic Decision-Making and a “Right to Explanation””, *AI Magazine*, vol. 38, núm. 3, 2017, pàg. 50-57.

GREEN, B., “The False Promise of Risk Assessments: Epistemic Reform and the Limits of Fairness”, *FAT* '20: Proceedings of the 2020 Conference on Fairness, Accountability, and Transparency*, 2020, pàg. 594-606.

- “The flaws of policies requiring human oversight of government algorithms”, *Computer Law and Security Review*, vol. 45, 2022, pàg. 1-22.

GREENLEAF, G., *The “Brussels Effect” of the EU’s “AI Act” on Data Privacy Outside Europe*, Privacy Laws and Business, International Report, núm. 171, Sydney, 2021.

GRGIĆ-HLAČA, N., REDMILES, E. M., GUMMADI, K. P. i WELLER, A., “Human perceptions of fairness in algorithmic decision making: A case study of criminal risk prediction”, *Proceedings of the 2018 World Wide Web Conference on World Wide Web*, 2018, pàg. 903-112.

GRUPO DE EXPERTOS DE ALTO NIVEL SOBRE INTELIGENCIA ARTIFICIAL, *Directrices éticas para una IA fiable*, Comisión Europea, Bruselas, 8 de abril de 2019.

GUASP DELGADO, J. i ARAGONESES ALONSO, P., *Derecho procesal civil*, Pamplona, Civitas, 2005.

GUIHOT, M. *et al.*, “Nudging Robots: Innovative Solutions to Regulate Artificial Intelligence”, *Vanderbilt Journal of Entertainment and Technology Law*, vol. 20, núm. 2, 2017, pàg. 385-456.

GUTIÉRREZ DE CABIEDES, P., *La Tutela jurisdiccional de los intereses supraindividuales: colectivos y difusos*, Elcano, Aranzadi, 1999.

HACKER, P., “Teaching fairness to artificial intelligence: existing and novel strategies against algorithmic discrimination under EU law”, *Common Market Law Review*, vol. 55, núm. 4, 2018, pàg. 1143-1185.

HACKER, P. i PASSOTH, J-H, “Varieties of AI Explanations under the Law. From GDPR to the AIA, and beyond”, a HOLZINGER, A., GOEBEL, R., FONG, R., MOON, T., MÜLLER, KR. i SAMEK, W. (eds.), *xxAI - Beyond Explainable AI*, Cham, Springer, 2022, pàg. 343-373.

HAKKARAINEN, J., “Naming something collective does not make it so: algorithmic discrimination and access to justice”, *Internet Policy Review. Journal on internet regulation*, vol. 10, núm. 4, 2021, pàg. 1-24.

HAMON, R, JUNKLEWITZ, H i SANCHEZ, I., *Robustness and Explainability of Artificial Intelligence*, JRC Technical Report, EU Science Hub, European Commission, Luxembourg, 2020.

HAN, B-C., *En el enjambre*, Barcelona, Herder, 2021.

HAVELKOVÁ, B., “Judicial Scepticism of Discrimination at the ECtHR”, a COLLINS, H., KHAITAN, T., (eds.), *Foundations of Indirect Discrimination Law*, Oxford, Hart Publishing, 2018, pàg. 83-104.

HEINRICHS, B., "Discrimination in the age of artificial intelligence", *AI & SOCIETY*, vol. 37, núm. 1, 2022, pàg. 143-154.

HELGUERO SÁINZ, J. i PARDO DE VERA, M. M., “Mecanismos de certificación en la nueva normativa de protección de datos de carácter personal (Comentario al artículo 42 RGPD y al artículo 39 LOPDGDD)”, a TRONCOSO REIGADA, A., (dir.) i VILLALBA CANO, L., (coord.), *Comentario al Reglamento General de Protección de*

Datos y a la Ley Orgánica de Protección de Datos Personales y Garantía de los Derechos Digitales, Tomo I, Pamplona, Aranzadi, 2021, pàg. 2421-2441.

HERNÁNDEZ, J.C., “Decisiones algorítmicas de perfilado: Régimen y garantías jurídicas”, *Revista Española de Derecho Administrativo*, núm. 203, 2020, pàg. 281-322.

HERRÁN ORTIZ, A. I., “Aproximación al derecho a la protección de datos personales en Europa. El Reglamento General de Protección de Datos personales a debate”, *Revista de Derecho, Empresa y Sociedad*, núm. 8, 2016, pàg. 179-200.

HIGH LEVEL GROUP ON NON-DISCRIMINATION, EQUALITY, AND DIVERSITY. SUBGROUP ON EQUALITY DATA, *Guidelines on improving the collection and use of equality data*, Brussels, 2018.

HILDEBRANDT, M. “The Dawn of a Critical Transparency Right for the Profiling Era” a BUS, J. *et al.* (eds.), *Digital Enlightenment Yearbook*, Amsterdam, IOS Press BV, 2012, pàg. 41-56.

HODGES, C., “Collective Redress: The Need for New Technologies”, *Journal of Consumer Policy*, vol. 42, núm. 1, 2019, pàg. 59-90.

HOFFMANN, A. L., “Where fairness fails: data, algorithms and the limits of antidiscrimination discourse”, *Information, Communication and Society*, vol. 22, núm. 7, 2019, pàg. 900-915.

HOFFMANN, H., VOGT, V., HAUER, M. P. i ZWEIG, K., “Fairness by awareness? On the inclusion of protected features in algorithmic decisions”, *Computer, Law and Security Review*, vol. 44, 2022, pàg. 1-12.

HOLSTEIN, K. *et al.* “Improving Fairness in Machine Learning Systems: What Do Industry Practitioners Need?”, *Proceedings of the ACM Conference on Human Factors in Computing Systems*, 2019, pàg. 1-16.

HOUSE OF LORDS SELECT COMMITTEE ON ARTIFICIAL INTELLIGENCE, *AI in the UK: Ready, willing and able*, London, 16 April 2018.

HUERGO LORA, A.J., “Una aproximación a los algoritmos desde el derecho administrativo”, a HUERGO LORA, A.J., (dir.) i DÍAZ GONZÁLEZ, G. M., (coord.),

La regulación de los algoritmos, Pamplona, Aranzadi, 2020, pàg. 23-87.

HUERTA ANGUIANO, J.A., ““Naturaleza Intrínseca”, “Contexto” o “Finalidad” en la determinación del carácter sensible de los datos personales”, *Estudios en derecho a la información*, núm. 10, 2020, pàg. 1-31.

INSTITUT MONTAIGNE, *Algorithmes: contrôle des biais S.V.P.*, Paris, Mars 2020.

IVANOVA, Y., “The Data Protection Impact Assessment as a Tool to Enforce Non-discriminatory AI”, *Privacy Technologies and Policy: 8th Annual Privacy Forum*, 2020, pàg. 3-24.

JABŁONOWSKA, A., KUZIEMSKI, M., NOWAK, A. M., MICKLITZ, H-W., PAŁKA, P. i SARTOR, G., *Consumer law and artificial intelligence. Challenges to the EU consumer law and policy stemming from the business’ use of artificial intelligence*, European University Institute (EUI), San Domenico di Fiesole, 2018.

JANSSEN, H. L., “Detecting New Approaches for a Fundamental Rights Impact Assessment to Automated Decision-Making”, *International Data Privacy Law*, vol. 10, núm. 1, 2020, pàg. 76-106.

JASANOFF, S., *Designs on Nature: Science and Democracy in Europe and the United States*, Princeton, New Jersey, Princeton University Press, 2011.

JIMÉNEZ-CASTELLANOS BALLESTEROS, I., "Decisiones automatizadas y transparencia administrativa: nuevos retos para los derechos fundamentales", *Revista Española de la Transparencia (RET)*, núm. 16, 2023, pàg. 191-215.

JOYCE, K. *et al.*, “Toward a Sociology of Artificial Intelligence: A Call for Research on Inequalities and Structural Change”, *Socius: Sociological Research for a Dynamic World*, vol. 7, 2021, pàg. 1-11.

KÁDÁR, T., “The legal standing of equality bodies”, *European equality law review*, núm. 1, 2019, pàg. 1-15.

KAMINSKI, M.E., “Binary Governance: Lessons from the GDPR’s Approach to Algorithmic Accountability”, *Southern California Law Review*, vol. 92, 2019, pàg. 1529-1616.

KAMINSKI, M. E. i MALGIERI, G., “Algorithmic impact assessments under the GDPR: producing multi-layered explanations”, *International Data Privacy Law*, vol. 11, núm. 2, 2021, pàg. 125-144.

KATELL, M., YOUNG, M., DAILEY, D., HERMAN, B., GUETLER, V., TAM, A., BINZ, C., RAZ, D., i KRAFFT, P., “Toward situated interventions for algorithmic equity: Lessons from the field”, *Proceedings of the 2020 Conference on Fairness, Accountability, and Transparency*, 2020, pàg. 45-55.

KATYAL, S. K., “Private Accountability in the Age of Artificial Intelligence”, *U.C.L.A. Law Review*, vol. 66, 2019, pàg. 54-141.

KEREN, H., “We insist! Freedom now: Does contract doctrine have anything constitutional to say?”, *Michigan Journal of Race Law*, vol. 11, núm. 1, 2005, pàg. 133-193.

KIM, H., “Cooperation and Coordination between the Financial Authorities: A Review of the Experiences of the United Kingdom, Norway, Sweden, and Korea”, *Seoul Journal of Economics*, vol. 22, núm. 3, 2009, pàg. 409-444.

KIM, P. T., “Auditing Algorithms for Discrimination”, *University of Pennsylvania Law Review Online*, vol. 166, 2017, pàg. 189-203.

KIRIMI, J.M. i MOTURI, C. A., “Application of data mining classification in employee performance prediction”, *International Journal of Computer Applications*, vol. 146, núm. 7, 2016, pàg. 28-35.

KLIMT, B. i YANG, Y., “Introducing the Enron Corpus”. *First Conference on Email and Anti-Spam (CEAS)*, 2004, pàg. 1-2.

KOENE, A., “Algorithmic Bias: Addressing Growing Concerns”, *IEEE Technology and Societal Magazine*, vol. 36, núm. 2, 2017, pàg. 31-32.

KOULU, R., “Human Control over Automation: EU Policy and AI Ethics”, *European Journal of Legal Studies*, vol. 12, núm. 1, 2020, pàg. 9-46.

KROLL, J. *et al.*, “Accountable algorithms”, *University of Pennsylvania Law Review*, vol. 165, núm. 3, 2017, pàg. 633-705.

KULLMANN, M., “Platform Work, Algorithmic Decision Making, and EU Gender Equality Law”, *International Journal of Comparative Labour Law and Industrial Relations*, vol. 34, núm. 1, 2018, pàg. 1-21.

LAHOTI, P., GUMMADI, K. P. i WEIKUM, G., “Learning individually fair data representations for algorithmic decision making”, *IEEE, 35th International Conference on Data Engineering (ICDE)*, 2019, pàg. 1334–1345.

- “Operationalizing individual fairness with pairwise fair representations”, *Proceedings of the VLDB Endowment*, vol.13, núm. 4, 2019, pàg. 506–518.

LEE, E., “Recognizing Rights in Real Time: The Role of Google in the EU Right To Be Forgotten”, *University of California, Davis Law Review*, vol. 49, núm. 3, 2016, pàg. 1017-1095.

LEESE, M., “The New Profiling: Algorithms, Black Boxes, and the Failure of AntiDiscriminatory Safeguards in the European Union”, *Security Dialogue*, vol. 45, núm. 5, 2014, pàg. 494-511.

LEPRI, B., STAIANO, J., SANGOKOYA, D., LETOUZÉ, E. i OLIVER, N., “The tyranny of data? the bright and dark sides of data-driven decision-making for social good”, a CERQUITELLI T., QUERCIA D., i PASQUALE F., (eds.), *Transparent Data Mining for Big and Small Data*, Switzerland, Springer, 2017, pàg. 3-22.

LERMAN, J., “Big Data and its exclusions”, *Stanford Law Review*, vol. 66, pàg. 55-63.

LOBACHEVA, A., i KASHTANOVA, E., “Social Discrimination in the epoch of artificial intelligence”, *Wisdom, Special Issue*, vol. 1, núm. 2, 2022, pàg. 97-105.

LÓPEZ PACHECO, J. A., “Movilización y acción colectiva por los derechos humanos en la paradoja de la institucionalización”, *Estudios Políticos (Universidad de Antioquia)*, vol. 51, 2017, pàg. 57-78.

LOWRY, S. i MACPHERSON, G., “A blot on the profession”, *British Medical Journal*, vol. 296, núm. 6623, 1988, pàg. 657-658.

LUM, K. i WILLIAM, I., “To predict and serve?”, *Significance*, vol.13, núm. 5, 2016, pàg. 14-19.

LUNDBERG, E., *Automated decision-making vs indirect discrimination. Solution or aggravation?*, Umeå University, Umeå, 2019.

LÜTZ, F., “Gender equality and artificial intelligence in Europe. Addressing direct and indirect impacts of algorithms on gender-based discrimination”, *ERA Forum*, vol. 23, núm. 1, 2022, pàg. 33-52.

MABBETT, D., “Aspirational Legalism and the Role of the Equality and Human Rights Commission in Equality Policy”, *The Political Quarterly*, vol 79, núm. 1, 2008, pàg. 45-52.

MACKINNON, C. A., *Feminism unmodified: Discourses on life and law*, Cambridge, Massachusetts, Harvard University Press, 1987.

- “Difference and Dominance: on sex discrimination”, a BARTLETT, K. T. i KENNEDY, R., (eds.), *Feminist legal theory: readings in law and gender*, Boulder, Westview Press, 1991, pàg. 81-94.

MAKKONEN, T., *Measuring Discrimination: Data Collection and EU Equality Law*, European Network of legal experts in the Non-discrimination field, Brussels, 2007.

MALGIERI, G. i COMANDÉ, G., “Sensitive-by-distance: quasi-health data in the algorithmic era”, *Information and Communications Technology Law*, vol. 26, núm. 3, 2017, pàg. 229-249.

MALISZEWSKA-NIENARTOWICZ, J., “Direct and Indirect Discrimination in the European Union Law. How to draw a Dividing Line?”, *International Journal of Social Sciences*, vol. 3, núm. 1, 2014, pàg. 41-55.

MANEIRO VÁZQUEZ, Y., *La discriminación por asociación: desafíos sustantivos y procesales. Desde su creación por el Tribunal de Justicia a su recepción por la doctrina judicial española*, Pamplona, Aranzadi, 2022.

MANN, M. i MATZNER, T., “Challenging algorithmic profiling: The limits of data protection and anti-discrimination in responding to emergent discrimination”, *Big Data and Society*, vol. 6, núm. 2, 2019, pàg. 1-11.

MANTELERO, A., “Personal data for decisional purposes in the age of analytics: From

an individual to a collective dimension of data protection”, *Computer Law & Security Review*, vol. 32, núm. 2, 2016, pàg. 239-255.

- “AI and Big Data: A Blueprint for a Human Rights, Social and Ethical Impact Assessment”, *Computer Law and Security Review*, vol. 34, núm. 4, 2018, pàg. 754-772.
- *Beyond Data. Human Rights, Ethical and Social Impact Assessment in AI*, The Hague, Asser Press, 2022,

MARCUS, M.P., SANTORINI, B. i MARCINKIEWICZ, M.A., “Building a Large Annotated Corpus of English: The Penn Treebank”, *Computational Linguistics, Special Issue on Using Large Corpora: II*, vol. 19, núm. 2, 1993, pàg. 313-330.

MARTÍNEZ MARTÍNEZ, R., “Inteligencia artificial desde el diseño. Retos y estrategias para el cumplimiento normativo”, *Revista Catalana de Dret Públic*, núm. 58, 2019, pàg. 64-81.

- “Cuestiones de ética jurídica al abordar proyectos de Big Data. El contexto del Reglamento General de Protección de Datos”, *Dilemata*, núm. 24, 2017, pàg. 151-164.

MARTÍNEZ GARCÍA, J. I., “El principio de igualdad y la producción de diferencias en el derecho”, a MINISTERIO DE JUSTÍCIA i CENTRO DE PUBLICACIONES, (eds.), *El principio de igualdad en la Constitución española: XI Jornadas de Estudio*, vol. 1, 1991, pàg. 539-562.

MERCAT-BURNS, M., “Systemic discrimination: Rethinking the tools of gender equality”, *European Equality Law Review*, núm. 2, 2018, pàg. 1-18.

MERRILL, R. A., “The Architecture of Government Regulation of Medical Products”, *Virginia Law Review*, vol. 82, núm. 8, 1996, pàg. 1753-1866.

MICHELAKAKI, C. i BARROS VALE, S., *Unlocking Data Protection By Design & By Default: Lessons from the Enforcement of Article 25 GDPR*, Future of Privacy Forum, Washington, DC, 19 May 2023.

MIRALLES LÓPEZ, R., “La evaluación de impacto relativa a la protección de datos

(Comentario al artículo 35 RGPD)”, a TRONCOSO REIGADA, A., (dir.) i VILLALBA CANO, L., (coord.), *Comentario al Reglamento General de Protección de Datos y a la Ley Orgánica de Protección de Datos Personales y Garantía de los Derechos Digitales*, Tomo I, Pamplona, Aranzadi, 2021, pàg. 2137-2162.

MONASTERIO ASTOBIZA, A., “Ética algorítmica: Implicaciones éticas de una sociedad cada vez más dominada por algoritmos”, *Dilemata*, núm. 24, 2017, pàg. 185-217.

MORENO CATENA, V., *Derecho Procesal Civil. Parte general*, Valencia, Tirant Lo Blanch, 2021.

MORENO MONTES DE OCA, B., “Autoridad de Control (Comentario al artículo 4.21 RGPD)”, a TRONCOSO REIGADA, A., (dir.) i VILLALBA CANO, L., (coord.), *Comentario al Reglamento General de Protección de Datos y a la Ley Orgánica de Protección de Datos Personales y Garantía de los Derechos Digitales*, Tomo I, Pamplona, Aranzadi, 2021, pàg. 795-805.

MULDER, J., *Indirect sex discrimination in employment. Theoretical analysis and reflections on the CJEU case law and national application of the concept of indirect sex discrimination*, European network of legal experts in gender equality and non-discrimination, Brussels, 2021.

MUÑOZ PAREDES, M.L., “«Big data» y contrato de seguro: los datos generados por los asegurados y su utilización por los aseguradores”, a HUERGO LORA, A.J., (dir.) i DÍAZ GONZÁLEZ, G. M., (coord.), *La regulación de los algoritmos*, Pamplona, Aranzadi, 2020, pàg. 129-162.

NICHOLSON PRICE II, W., “Regulating Black-Box Medicine”, *Michigan Law Review*, vol. 116, núm. 3, 2017, pàg. 421-474.

NIETO MORALES, C. i ABASCAL MONEDERO, P. J., *Reclamaciones en Materia de Consumo*, Madrid, Dykinson, 2016.

NIEVA FENOLL, J., *Inteligencia artificial y proceso judicial*, Madrid, Marcial Pons, 2018.

NISSENBAUM, H., *Privacy in Context: Technology, Policy, and the Integrity of Social Life*, Stanford, Stanford University Press, 2009.

NOBLE, S. U., *Algorithms of Oppression: How Search Engines Reinforce Racism*, New York, New York University Press, 2018.

NOTO LA DIEGA, G., “Against the Dehumanisation of Decision-Making Algorithmic Decisions at the Crossroads of Intellectual Property, Data Protection, and Freedom of Information”, *JIPITEC- Journal of Intellectual Property, Information Technology and E-Commerce Law*, vol. 9, núm. 3, 2018, pàg. 3-34.

OBASOGIE, O. K., “The return of Biological Race? Regulating Race and Genetics through Administrative Agency Race Impact Assessments”, *Southern California Interdisciplinary Law Journal*, vol. 22, núm. 1, 2012, pàg 1-64.

OBSERVATORI D'ÈTICA EN INTEL·LIGÈNCIA ARTIFICIAL DE CATALUNYA (OEIAC), *Intel·ligència artificial, ètica i societat: Una mirada i discussió a través de la literatura especialitzada i d'opinions expertes*, Girona, 2021.

OCDE, *Directrices sobre protección de la privacidad y flujos transfronterizos de datos personales*, Paris, 23 de septiembre de 1980.

OECD, *Profiling tools for early identification of jobseekers who need extra suport*, Paris, 2018.

OKUN, A. M., *Equality and Efficiency: The Big Tradeoff*, Washington, D.C., Brookings Institution Press, 1975.

O'NEIL, C., *Weapons of Math Destruction. How big data increases inequality and threatens democracy*, New York, Broadway Books, 2017.

ORGANISATION FOR SECURITY AND CO-OPERATION IN EUROPE (OSCE), *Use of actio popularis in cases of discrimination*, Skopje, November 2016.

ORTELLS RAMOS, M., “Tutela judicial civil colectiva y nuevos modelos de los servicios de defensa jurídica en España”, *Revista General de Derecho Procesal*, vol. 48, 2019, pàg. 1-46.

PÁEZ, A., “Negligent Algorithmic Discrimination”, *Law and Contemporary Problems*, vol. 84, núm. 3, 2021, pàg. 19-33.

PALMA ORTIGOSA, A., “Decisiones Automatizadas en el RGPD. El uso de algoritmos en el contexto de la protección de datos” *Revista General de Derecho Administrativo*, núm. 50, 2019, pàg. 1-35.

- *Decisiones automatizadas y protección de datos. Especial atención a los sistemas de inteligencia artificial*, Madrid, Dykinson, 2022.

PARLAMENTO EUROPEO, *Resolución del Parlamento Europeo, de 14 de marzo de 2017, sobre las implicaciones de los macrodatos en los derechos fundamentales: privacidad, protección de datos, no discriminación, seguridad y aplicación de la ley*, 2016/2225(INI), Bruselas, 2017.

- *Resolución del Parlamento Europeo sobre «Hacia un planteamiento europeo coherente del recurso colectivo»*, 2011/2089(INI), Bruselas, 2 de febrero de 2012.
- *Resolución del Parlamento Europeo, de 21 de enero de 2021, sobre la estrategia de la Unión para la igualdad de género*, 2019/2169(INI), Bruselas, 2021.

PARRILLA CASTILLO, J.A., “La Elaboración de Perfiles Políticos en España tras la Sentencia del Tribunal Constitucional de 22 de mayo de 2019 por la que se declara Inconstitucional el Art. 58 bis.1 de la Ley Orgánica del Régimen Electoral General - Perfilado político en España tras el pronunciamiento del Tribunal Constitucional de España de 22 de mayo de 2019”, *Anuario do NOVA Consumer Lab*, Año 1, 2019, pàg. 67-80.

PASQUAL HUERTA, P., “El derecho de rectificación (Comentario al artículo 16 RGPD)”, a TRONCOSO REIGADA, A., (dir.) i VILLALBA CANO, L., (coord.), *Comentario al Reglamento General de Protección de Datos y a la Ley Orgánica de Protección de Datos Personales y Garantía de los Derechos Digitales*, Tomo I, Pamplona, Aranzadi, 2021, pàg. 1529-1556.

- “El derecho a la limitación de tratamiento (Comentario al artículo 18 RGPD)”, a TRONCOSO REIGADA, A., (dir.) i VILLALBA CANO, L., (coord.), *Comentario al Reglamento General de Protección de Datos y a la Ley Orgánica de Protección de Datos Personales y Garantía de los Derechos Digitales*, Tomo I, Pamplona, Aranzadi, 2021, pàg. 1605-1632.

PASQUALE, F., *The Black box society: The secret algorithms that control Money and information*, Cambridge (Massachusetts), Harvard University Press, 2015.

PERETE RAMÍREZ, C. i GARCÍA MEXÍA, P., “La propiedad sobre el dato. ¿Cabe una vertiente patrimonial de la protección de datos?”, a TRONCOSO REIGADA, A., (dir.) i VILLALBA CANO, L., (coord.), *Comentario al Reglamento General de Protección de Datos y a la Ley Orgánica de Protección de Datos Personales y Garantía de los Derechos Digitales*, Tomo I, Pamplona, Aranzadi, 2021, pàg. 1681-1700.

PLAZA PENADÉS, J., “Aspectos legales del Big Data y la Inteligencia Artificial”, a PEDREÑO MUÑOZ, A., MORENO IZQUIERDO, L., PLAZA PENEDÉS, J. i IRANZO CABRERA, M., *Big Data e Inteligencia Artificial: una visión económica y legal de estas tecnologías disruptives*, Fundació Parc Científic Universitat de València, 2019, pàg. 28-43.

POLITOU, E. *et al.*, “Backups and the right to be forgotten in the GDPR: An uneasy relationship”, *Computer Law & Security Review: The International Journal of Technology Law and Practice*, vol. 34, núm. 6, 2018, pàg. 1247-1257.

POLO ROCA, A., “Datos, datos, datos: el dato personal, el dato no personal, el dato personal compuesto, la anonimización, la pertenencia del dato y otras cuestiones sobre datos”, *Estudios de Deusto*, vol. 69, núm. 1, 2021, pàg. 211-240.

PONCE SOLÉ, J., "Las relaciones entre inteligencia artificial, regulación y ética, con especial atención al sector público", *Revista General de Derecho Administrativo*, núm. 61, 2022, pàg. 1-29.

POULLET, Y. i DINANT, J-M., *Report on the application of data protection principles to the worldwide telecommunication network*, Consultive committee of the Convention for the protection of individuals with regard to automatic processing of personal data, T-PD (2004) 04 final, Strasbourg, 12 December 2004.

PRIETO HERGUETA, J., “Los códigos de conducta (Comentario a los artículos 40 y 41 RGPD y a los artículos 38 y Disposición transitoria segunda LOPDGDD)”, a TRONCOSO REIGADA, A., (dir.) i VILLALBA CANO, L., (coord.), *Comentario al Reglamento General de Protección de Datos y a la Ley Orgánica de Protección de Datos Personales y Garantía de los Derechos Digitales*, Tomo I, Pamplona, Aranzadi, 2021, pàg. 2357-2377.

PUIG HERNÁNDEZ, M. A., *La protección de la libertad ideológica frente al trato discriminatorio en la relación jurídico-laboral*, Tesi doctoral, Universitat de Barcelona, Barcelona, 2018.

PURTOVA, N., “Do property rights in personal data make sense after the big data turn? Individual control and transparency”, *Journal of Law and Economic Regulation*, vol. 10, núm. 2, 2017, pàg. 64-78.

- “The law of everything: Broad concept of personal data and future of EU data protection law”, *Law, Innovation and Technology*, vol. 10, núm. 1, 2018, pàg. 40–81.

RAAB, C. D., “Information privacy, impact assessment, and the place of ethics”, *Computer Law and Security Review*, vol. 37, 2020, pàg. 1-16.

RANCHORDÁS, S. i SCHUURMANS, Y., “Outsourcing the welfare state: the role of private actors in welfare fraud investigations”, *European Journal of Comparative Law and Governance*, vol. 7, núm. 1, 2020, pàg. 5-42.

REBOLLO DELGADO, L., “Las condiciones de tratamiento de categorías especiales de datos (Comentario al artículo 9 RGPD y artículo 9 LOPDGDD)”, a TRONCOSO REIGADA, A., (dir.) i VILLALBA CANO, L., (coord.), *Comentario al Reglamento General de Protección de Datos y a la Ley Orgánica de Protección de Datos Personales y Garantía de los Derechos Digitales*, Tomo I, Pamplona, Aranzadi, 2021, pàg. 1013-1040.

REISMAN, D., SCHULTZ, J., CRAWFORD, K. i WHITTAKER, M., *Algorithmic Impact Assessments: a Practical Framework for Public Agency Accountability*, AI Now Institute, New York, April 2018.

REQUENA CASANOVA, M., “La tutela judicial del principio general de igualdad de trato en la Unión Europea: una jurisprudencia expansiva basada en una jerarquía de motivos discriminatorios”, *Revista de Derecho Comunitario Europeo*, núm. 40, pàg. 767-793.

REY MARTÍNEZ, F., “La discriminación múltiple, una realidad antigua, un concepto nuevo”, *Revista española de derecho constitucional*, núm. 84, 2008, pàg. 251-283.

- *Derecho Antidiscriminatorio*, Navarra, Aranzadi, 2019.
- “O organismo español para a igualdade: claves para a súa converxencia co modelo europeo”, *Revista Galega de Administración Pública*, núm 61, 2021, pàg. 311-342.

RICHARDS, N. M. i KING, J. H., “Three Paradoxes of Big Data”, *Stanford Law Review Online*, vol. 66, 2013, pàg. 41-46.

RINGELHEIM, J., “The Burden of Proof in Antidiscrimination Proceedings. A Focus on Belgium, France and Ireland”, *European Equality Law Review*, núm. 2, 2019, pàg. 49-64.

RIVAS VALLEJO, P., *La aplicación de la Inteligencia Artificial del trabajo y su impacto discriminatorio*, Navarra, Aranzadi, 2020.

- “Un matiz nuevo en la delimitación del ámbito subjetivo de protección de las directivas antidiscriminación: diferencia de trato entre sujetos de la misma categoría protegida”, *Revista de Jurisprudencia Laboral*, núm. 1/2021, 2021, pàg. 1-11.
- “Discriminación algorítmica: detección, prevención y tutela”, *XXXI Jornades Catalanes de Dret Social (“Trellal, discriminació i COVID”)*, 2021, pàg. 1-79.
- “Sesgos de género en el uso de inteligencia artificial para la gestión de las relaciones laborales: análisis desde el derecho antidiscriminatorio”, *e-Revista Internacional de la Protección Social (e-RIPS)*, vol. VII, núm. 1, 2022, pàg. 52-83.

- “Análisis desde el derecho antidiscriminatorio”, a RIVAS VALLEJO, P., (dir.), *Discriminación algorítmica en el ámbito laboral: perspectiva de género e intervención*, Navarra, Aranzadi, 2022, pàg. 415-476.

ROARK, M.L., “Human Impact Statements”, *Washburn Law Journal*, vol. 54, núm. 3, 2015, pàg. 649-682.

ROBERTS, H. *et al.*, “Artificial intelligence regulation in the United Kingdom: a path to good governance and global leadership?”, *Internet Policy Review*, vol. 12, núm. 2, 2023, pàg. 1-31.

ROBERTS, J.L., “Protecting Privacy to prevent discrimination”, *William and Mary Law Review*, vol. 56, núm. 6, 2015, pàg. 2097-2174.

RODRÍGUEZ-PIÑERO y BRAVO-FERRER, M., “Los contornos de la discriminación”, *Temas Laborales*, núm. 162, 2022, pàg. 11-18.

ROIG, A., “Safeguards for the Right Not To Be Subject to a Decision Based Solely on Automated Processing (Article 22 GDPR)”, *European Journal of Law and Technology*, vol. 8, núm. 3, 2017, pàg. 1-17.

- *Las Garantías frente a las decisiones automatizadas. Del Reglamento General de Protección de Datos a la gobernanza algorítmica*, Barcelona, Bosch, 2020.
- “Límits del principi de privadesa pel disseny”, *Revista Catalana de Dret Públic*, núm. 64, 2022, pàg. 174-186.

RORIVE, I., “A comparative and European examination of National Institutions in the field of racism and discrimination”, a BOYLE, K., (ed.), *New Institutions for human Rights Protection*, Oxford, Oxford University Press, 2009, pàg. 137-173.

RUBIO RIVAS, E., “Régimen jurídico de los procedimientos tramitados por la Agencia Española de Protección de Datos en caso de posible vulneración de la normativa de protección de datos personales (Comentario al artículo 63 LOPDGDD)”, a TRONCOSO REIGADA, A., (dir.) i VILLALBA CANO, L., (coord.), *Comentario al Reglamento General de Protección de Datos y a la Ley Orgánica de Protección de Datos Personales y Garantía de los Derechos Digitales*, Tomo II, Pamplona, Aranzadi, 2021, pàg. 2699-

2729.

RUDIN, C., “Stop explaining black box machine learning models for high stakes decisions and use interpretable models instead”, *Nature Machine Intelligence*, vol.1, núm. 5, 2019, pàg. 206-215.

SÁEZ LARA, C., “El algoritmo como protagonista de la relación laboral. Un análisis desde la perspectiva de la prohibición de discriminación”, *Temas Laborales: Revista andaluza de Trabajo y bienestar social*, núm. 155, 2020, pàg. 41-60.

SÁNCHEZ-MONEDERO, J., DENCİK, L., i EDWARDS, L., “What Does It Mean to “solve” the Problem of Discrimination in Hiring?: Social, Technical and Legal Perspectives from the UK on Automated Hiring Systems”, *FAT* '20: Proceedings of the 2020 Conference on Fairness, Accountability, and Transparency*, 2020, pàg. 458-568.

SANCHO LÓPEZ, M., “Estrategias legales para garantizar los derechos fundamentales frente a los desafíos del Big Data”, *Revista General de Derecho Administrativo*, núm. 50, 2019, pàg. 1-28.

SANTAMARÍA RAMOS, F. J., “El derecho a la portabilidad: Un nuevo derecho esencial en materia de protección de datos (Comentario al artículo 20 RGPD y al artículo 17 LOPDGDD)”, a TRONCOSO REIGADA, A., (dir.) i VILLALBA CANO, L., (coord.), *Comentario al Reglamento General de Protección de Datos y a la Ley Orgánica de Protección de Datos Personales y Garantía de los Derechos Digitales*, Tomo I, Pamplona, Aranzadi, 2021, pàg. 1655-1680.

- “La designación del delegado de protección de datos (Comentario al artículo 37 RGPD y al artículo 34 LOPDGDD)”, a TRONCOSO REIGADA, A., (dir.) i VILLALBA CANO, L., (coord.), *Comentario al Reglamento General de Protección de Datos y a la Ley Orgánica de Protección de Datos Personales y Garantía de los Derechos Digitales*, Tomo I, Pamplona, Aranzadi, 2021, pàg. 2237-2265.

SCHERER, M. U., “Regulating Artificial Intelligence Systems: Risks, Challenges, Competencies, and Strategies”, *Harvard Journal of Law & Technology*, vol. 29, núm. 2, 2016, pàg. 353-400.

SCHEUERMAN, M.K., WADE, K., LUSTING, C. i BRUBAKER, J.R., “How We’ve Taught Algorithms to See Identity: Constructing Race and Gender in Image Databases for Facial Analysis”, *Proceedings of the ACM on Human-Computer Interaction*, vol. 4, article 58, 2020, pàg. 1-35.

SCHREURS, W., “Cogitas, Ergo Sum. The role of data protection Law and non-discrimination law in group profiling in the private sector”, a HIDELEBRANDT, M. i GUTWIRTH, S., (eds.), *Profiling the European citizen: Cross-Disciplinary Perspectives*, Berlin, Springer Dordrecht, 2008, pàg. 241-270.

SEBASTIANI, L., “Investigando los límites de la lucha legal contra el racismo: El marco español de antidiscriminación por origen racial o étnico”, *Oñati Socio-legal series*, vol. 11, núm. 3, 2021, pàg. 753-786.

SELBST, A. D., “Disparate Impact in Big Data Policing”, *Georgia Law Review*, vol. 52, núm. 1, 2017, pàg. 109-195.

- “An institutional view of Algorithmic Impact Assessments”, *Harvard Journal of Law and Technology*, vol. 35, núm.1, 2021, pàg. 117-191.

SELBST, A. D. i BAROCAS, S., “The Intuitive Appeal of Explainable Machines”, *Fordham Law Review*, vol. 87, núm. 3, 2018, pàg. 1085-1139.

SEMPERE SAMANIEGO, J., “Organismo de certificación (Comentario al artículo 43 RGPD y al artículo 39 LOPDGDD)”, a a TRONCOSO REIGADA, A., (dir.) i VILLALBA CANO, L., (coord.), *Comentario al Reglamento General de Protección de Datos y a la Ley Orgánica de Protección de Datos Personales y Garantía de los Derechos Digitales*, Tomo I, Pamplona, Aranzadi, 2021, pàg. 2443-2456.

SERRANO PÉREZ, M. M., “Los códigos de conducta (Comentario artículo 40 RGPD y al artículo 38 y Disposición transitoria segunda LOPDGDD)”, a TRONCOSO REIGADA, A., (dir.) i VILLALBA CANO, L., (coord.), *Comentario al Reglamento General de Protección de Datos y a la Ley Orgánica de Protección de Datos Personales y Garantía de los Derechos Digitales*, Tomo I, Pamplona, Aranzadi, 2021, pàg. 2379-2407.

SIMÓN CASTELLANO, P., “Allende una teoría general de las garantías jurídicas para

una inteligencia artificial confiable”, a BALAGUER CALLEJÓN, F. I COTINO HUESO, L. (coords), *Derecho Público de la Inteligencia Artificial*, Zaragoza, Fundación Manuel Giménez Abad, 2023, pàg. 111-148.

SMUHA, N. A., “Beyond the individual: governing AI’s societal harm”, *Internet Policy Review*, vol. 10, núm. 3, 2021, pàg. 1-32.

SOLOW-NIEDERMAN, A., “Administering Artificial Intelligence”, *Southern California Law Review*, vol. 93, 2020, pàg. 633-696.

SORIANO ARNANZ, A., “Decisiones automatizadas y discriminación: aproximación y propuestas generales”, *Revista General de Derecho Administrativo*, núm. 56, 2021, pàg. 1-45.

- *Data Protection for the prevention of algorithmic discrimination. Protecting from discrimination and other harms caused by algorithms through privacy in the EU: possibilities, shortcomings and proposals*, Pamplona, Aranzadi, 2021.
- “La propuesta de Reglamento de Inteligencia Artificial de la UE y los sistemas de alto riesgo”, *Revista General de Derecho de los Sectores Regulados*, núm. 8, 2021, pàg. 1-24.
- “Decisiones automatizadas: problemas y soluciones jurídicas. Más allá de la protección de datos”, *Revista de Derecho Público: Teoría y Método*, vol. 3, 2021, pàg. 85-127.
- “La aplicación del marco jurídico europeo en materia de igualdad y no discriminación al uso de aplicaciones de Inteligencia Artificial”, a BONORINO RAMÍREZ, P. R., FERNÁNDEZ ACEVEDO, R. i VALCÁRCEL FERNÁNDEZ, P., (dirs.), *Nuevas normatividades: inteligencia artificial, derecho y genero*, Navarra, Aranzadi, 2021, pàg. 63-88.
- “Discriminación algorítmica: garantías y protección jurídica”, a COTINO HUESO, L., (dir.) i BAUZA REILLY, M., (coord.), *Derechos y garantías ante la inteligencia artificial y las decisiones automatizadas*, Navarra, Aranzadi, 2022, pàg. 139-170.

- “Creating non-discriminatory Artificial Intelligence systems: balancing the tensions between code granularity and the general nature of legal rules”, *Revista d’Internet, Dret i Política*, núm. 38, 2023, pàg. 1-12.

SPRINGER, A., GARCIA-GATHRIGHT, J. i CRAMER, H., “Assessing and Addressing Algorithmic Bias– But Before We Get There”, *AAAI 2018 Spring Symposium Series*, 2018, pàg. 450-454.

SQUIRE, J., “Intersecting inequalities”, *International Feminist Journal of politics*, vol. 11, núm. 4, 2009, pàg. 496-512.

STRAK, L., “Facial Recognition is the Plutonium of AI”, *XRDS: Crossroads, The ACM Magazine for Students*, vol. 25, núm. 3, 2019, pàg. 50-55.

SURESH, H. i GUTTAG, J., “Framework for Understanding Unintended Sources of Harm Throughout the Machine Learning Life Cycle”, *EAAMO 2021: ACM Conference on Equity and Access in Algorithms, Mechanisms, and Optimization*, 2021, pàg. 4-6.

SWEENEY, L., “Discrimination in Online Ad Delivery”, *Acm Queue*, vol. 11, núm. 3, 2013, pàg. 10-29.

TAMÒ-LARRIEUX, A., *Designing for Privacy and Its Legal Framework: Data Protection by Design and Default for the Internet of Things*, Cham, Springer, 2018.

THE ROYAL SOCIETY, *Machine learning: the power and promise of computers that learn by example*, London, April 2017.

TOBLER, C., *Limits and potential of the concept of indirect discrimination*, European Network of Legal Experts in Anti-Discrimination, Brussels, 2008.

- *Indirect discrimination under Directives 2000/43 and 2000/78*, European network of legal experts in gender equality and non-discrimination, Brussels, 2022.

TORRALBA, F., *L’ètica algorítmica*, Barcelona, Edicions 62, 2022.

TOSONI, L., “The right to object to automated individual decisions: Resolving the ambiguity of Article 22(1) of the General Data Protection Regulation”, *International Data*

Privacy Law, vol. 11, núm. 2, 2021, pàg. 145-162.

TUTT, A., “An FDA for algorithms”, *Administrative Law Review*, vol. 69, núm. 1, 2017, pàg. 83-123.

UNITED KINGDOM INFORMATION COMMISSIONER’S OFFICE, *Feedback request profiling and decision making*, London, 6 April 2017.

VAN BEKKUM, M. i ZUIDERVEEN, F., “Using sensitive data to prevent discrimination by artificial intelligence: Does the GDPR need a new exception?” *Computer Law and Security Review*, vol. 48, 2023, pàg. 1-12.

VAN DEN BRINK, M. i DUNNE, P., *Trans and intersex equality rights in Europe - a comparative analysis*, European network of legal experts in gender equality and non-discrimination, Brussels, 2018.

VAN GESTEL, R. i MICKLITZ, H.W., “European integration through standardization: How judicial review is breaking down the club house of private standardization bodies”, *Common market law review*, vol 50, núm. 1, 2013, pàg. 145-181.

VAN ZOONEN, L., “Performance and Participation in the Panopticon: Instruments for Civic Engagement with Urban Surveillance Technologies”, a JACOBS, G., SUOJANEN, I., HORTON, K. i BAYERL, P., (eds.), *International Security Management. Advanced Sciences and Technologies for Security Applications*, Cham, Springer, 2021, pàg. 243-254.

VARGAS CELIX, M. V., “Entre la confianza ciudadana y el impulso tecnológico. Retos jurídicos de la Inteligencia Artificial a la luz de la propuesta de Reglamento de la UE”, *IUS ET SCIENTIA*, vol. 8, núm. 1, 2022, pàg. 176-188.

VEALE, M. i BINNS, R., “Fairer machine learning in the realworld: Mitigating discrimination without collecting sensitive data”, *Big Data & Society*, vol. 4, núm. 2, 2017, pàg. 1-17.

VEALE, M., BINNS, R., i AUSLOOS, J. “When Data Protection by Design and Data Subject Rights Clash”, *International Data Privacy Law*, vol. 8, núm. 2, 2018, pàg. 105-123.

VEALE, M. i EDWARDS, L., “Slave to the Algorithm? Why a 'Right to an Explanation' Is Probably Not the Remedy You Are Looking For”, *Duke Law and Technology Review*, vol. 16, núm. 1, 2017, pàg. 18-84.

- “Clarity, surprises, and further questions in the Article 29 Working Party draft guidance on automated decision-making and profiling”, *Computer law and security review*, vol. 34, núm. 2, 2018, pàg. 398-404.

VEALE, M. i ZUIDERVEEN, F., “Demystifying the Draft EU Artificial Intelligence Act. Analysing the good, the bad, and the unclear elements of the proposed approach”, *Computer Law Review International*, vol. 22, núm. 4, 2021, pàg. 97-112.

VESTRI, G., "La inteligencia artificial ante el desafío de la transparencia algorítmica. una aproximación desde la perspectiva jurídico-administrativa", *Revista Aragonesa de Administración Pública*, núm. 56, 2021, pàg. 368-398.

VICENTE PALACIO, A., "La inteligencia artificial en la ley integral para la igualdad y no discriminación", *Revista General de Derecho del Trabajo y de la Seguridad Social*, núm. 64, 2023, pàg. 125-150.

VILLALBA CANO, L., “La representación de los interesados (Comentario al artículo 80 RGPD)”, a TRONCOSO REIGADA, A., (dir.) i VILLALBA CANO, L., (coord.), *Comentario al Reglamento General de Protección de Datos y a la Ley Orgánica de Protección de Datos Personales y Garantía de los Derechos Digitales*, Tomo II, Pamplona, Aranzadi, 2021, pàg. 3019-3040.

- “El derecho a presentar una reclamación ante una autoridad de control (Comentario al artículo 77 RGPD)”, a TRONCOSO REIGADA, A., (dir.) i VILLALBA CANO, L., (coord.), *Comentario al Reglamento General de Protección de Datos y a la Ley Orgánica de Protección de Datos Personales y Garantía de los Derechos Digitales*, Tomo II, Pamplona, Aranzadi, 2021, pàg. 2967-2987.
- “El derecho a la tutela judicial efectiva contra una autoridad de control (Comentario al artículo 78 RGPD)”, a TRONCOSO REIGADA, A., (dir.) i VILLALBA CANO, L., (coord.), *Comentario al Reglamento General de Protección de Datos y a la Ley Orgánica de Protección de Datos Personales y Garantía de los Derechos Digitales*, Tomo II, Pamplona, Aranzadi, 2021, pàg. 2988-3018.

y *Garantía de los Derechos Digitales*, Tomo II, Pamplona, Aranzadi, 2021, pàg. 2989-3005.

- “El derecho a la tutela judicial efectiva contra un responsable o encargado del tratamiento (Comentario al artículo 79 RGPD)”, a TRONCOSO REIGADA, A., (dir.) i VILLALBA CANO, L., (coord.), *Comentario al Reglamento General de Protección de Datos y a la Ley Orgánica de Protección de Datos Personales y Garantía de los Derechos Digitales*, Tomo II, Pamplona, Aranzadi, 2021, pàg. 3007-3018.

VILLANI, C., *Donner un sens à l'intelligence artificielle: pour une stratégie nationale et européenne*, Mission confiée par le Premier Ministre Édouard Philippe, Paris, 2018.

WACHTER, S., “Affinity profiling and discrimination by association in online behavioural advertising”, *Berkeley Technology Law Journal*, vol. 35, núm. 2, 2020, pàg. 367-430.

WACHTER, S. i MITTELSTADT, B., “A Right to Reasonable Inferences: Re-Thinking Data Protection Law in the Age of Big Data and AI”, *Columbia Business Law Review*, núm. 2, 2019, pàg. 494-620.

WACHTER, S., MITTELSTADT, B. i FLORIDI, L., “Why a Right to Explanation of Automated Decision-Making Does Not Exist in the General Data Protection Regulation”, *International Data Privacy Law*, vol.7, núm. 2, 2017, pàg. 76-99.

WACHTER, S., MITTELSTADT, B., i RUSSEL, C., “Counterfactual explanations without opening the black box: Automated decisions and the GDPR”, *Harvard Journal of Law & Technology*, vol. 31, núm. 2, 2018, pàg. 841-887.

- “Why fairness cannot be automated: Bridging the gap between EU non-discrimination law and AI”, *Computer Law and Security Review*, vol 41, 2021, pàg. 1-72.

WAGNER, B., “Liable, but Not in Control? Ensuring Meaningful Human Agency in Automated Decision-Making Systems”, *Policy & Internet*, vol. 11, núm. 1, 2019, pàg. 104–122.

WANG, Y. i KOSINSKI, M., “DeepNeural Networks Are More Accurate Than Humans at Detecting Sexual Orientation from Facial Images” *Journal of Personality and Social Psychology*, vol. 114, núm. 2, 2018, pàg. 246-257.

WHITTAKER, M. *et al.*, *Disability, Bias and AI*, AI Now Institute, New York, November 2019.

WIERINGA, M., ““Hey SyRI, tell me about algorithmic accountability”: Lessons from a landmark case”, *Data and Policy*, vol. 5, núm. 2, 2023, pàg. 1-24.

WILLIAMS, B. A., BROOKS, C. F. i SHAMARGAD, Y., “How Algorithms Discriminate Based on Data they Lack: Challenges, Solutions, and Policy Implications”, *Journal of Information Policy*, vol. 8, 2018, pàg. 78-115.

WRIGHT, D. i FRIEDEWALD, M., “Integrating Privacy and Ethical Impact Assessments”, *Science and Public Policy*, vol. 40, núm. 6, 2013, pàg. 755-766.

WRIGHT, D. i RAAB, C.D., “Constructing a Surveillance Impact Assessment”, *Computer Law and Security Review*, vol. 28, núm. 6, 2012, pàg. 613-626.

WOLTERS, P.T.J., “The Control by and Rights of the Data Subject Under the GDPR”, *Journal of Internet Law*, vol. 22, núm. 1, 2018, pàg. 7-18.

XENIDIS, R., *Shaking the normative foundations of eu equality law: evolution and hierarchy between market integration and human rights rationales*, Working Paper, EUI LAW, 2017/04, European Regulatory Private Law Project (ERPL), San Domenico di Fiesole, 2017.

- “Turning EU equality law to algorithmic discrimination: Three pathways to resilience”, *Maastricht Journal of European and Comparative Law*, vol. 27, núm. 6, 2020, pàg. 736-758.

XENIDIS, R. i SENDEN, L., “EU Non-discrimination law in the era of artificial intelligence: mapping the challenges of algorithmic discrimination”, a BERNITZ, U., GROUSSOUT, X., PAJU, J. i DE VRIES, S., (eds.), *General Principles of EU Law and the EU Digital Order*, Alphen aan den Rijn, Kluwer Law International, 2020, pàg. 151-182.

YANISKY-RAVID, S. i HALLISEY, S., “Equality and privacy by design: a new model of artificial intelligence data transparency via auditing, certification and safe harbor regimes”, *Fordham Urban Law Journal*, vol. 46, núm. 2, 2019, pàg. 428-486.

ZARSKY, T., “An analytic challenge: discrimination theory in the age of predictive analytics”, *A Journal of Law and Policy for the Information Society*, vol. 14, núm. 1, 2017, pàg. 11-35.

ZARSKY, T. Z., “Incompatible: The GDPR in the Age of Big Data”, *Seton Hall Law Review*, vol. 47, núm. 4, 2017, pàg. 995-1020.

ZEHLIKE, M., HACKER, P., i WIEDEMANN, E., “Matching Code and Law: Achieving algorithmic fairness with optimal transport”, *Data Mining and Knowledge Discovery*, vol. 34, núm. 1, 2020, 163–200.

ŽLIOBAITĖ, I., CUSTS, B., “Using sensitive personal data may be necessary for avoiding discrimination in data-driven decision models”, *Artificial Intelligence Law*, vol. 24, núm. 2, 2016, pàg. 183–201.

ZUDDAS, P., “Decisión algorítmica y discriminación”, *Revista General de Derecho Administrativo*, núm. 60, 2022, pàg. 1-13.

ZUIDERVEEN, F., *Discrimination, artificial intelligence, and algorithmic decision-making*, Council of Europe, Strasbourg, 2018.

- “Strengthening legal protection against discrimination by algorithms and artificial intelligence”, *The International Journal of Human Rights*, vol. 24, núm. 10, 2020, pàg. 1572-1593.

ZUIDERVEEN, F. J. i POORT, J., “Online price discrimination and EU data privacy law”, *Journal of Consumer Policy*, vol. 40, 2017, pàg. 347-366.

BIBLIOGRAFIA WEB

Artificial Intelligence. Britannica. 2023. Disponible a: <https://www.britannica.com/technology/artificial-intelligence>. Darrera visita el 2 de setembre de 2023.

ANGWIN, J. i PARRIS Jr., T., “Facebook lets advertisers exclude users by race”, ProPublica, 28 d’octubre de 2016. Disponible a: <https://www.propublica.org/article/facebook-lets-advertisers-exclude-users-by-race>. Darrera visita el 2 de setembre de 2023.

ANGWIN, J. *et al.*, “Machine bias: There’s software used across the country to predict future criminals. And it’s biased against blacks”, ProPublica, 23 de maig de 2016. Disponible a: <https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing>. Darrera visita el 2 de setembre de 2023.

BELLAMY, R. K. E. *et al.*, “AI Fairness 360: An Extensible Toolkit for Detecting, Understanding, and Mitigating Unwanted Algorithmic Bias”, 2018, pàg. 1-20. Disponible a: <https://arxiv.org/>. Darrera visita el 2 de setembre de 2023.

CECCHI-DIMEGLIO, P., “How gender bias corrupts performance reviews, and what to do about it”, *Harvard Business Review*, 12 d’abril de 2017. Disponible a: <https://hbr.org/2017/04/how-gender-bias-corrupts-performance-reviews-and-what-to-do-about-it>. Darrera visita el 2 de setembre de 2023.

CONSEJO DE TRANSPARENCIA Y BUEN GOBIERNO, Contestación a la demanda en el Procedimiento Ordinario 18/2019 seguido en el Juzgado Central Contencioso-Administrativo núm. 8, de 5 de septiembre de 2019. Disponible a: <https://civio.app.box.com/s/q7b42raq1xxsrjhg8owddnbjm8yn6ga> Darrera visita el 2 de setembre de 2023.

CRAWFORD, K., “Think again: Big Data”, *Foreign Policy Review*, 10 de maig de 2013. Disponible a: <https://foreignpolicy.com/2013/05/10/think-again-big-data/>. Darrera visita el 2 de setembre de 2023.

DIAKOPOULOS, N. *et al.*, “Principles for Accountable Algorithms and a Social Impact Statement for Algorithms”, 2016. Disponible a: <http://sorelle.friedler.net/papers/principles.pdf>. Darrera visita el 2 de setembre de 2023.

Discriminació. TERMCAT Centre de Terminologia. 2023. Disponible a <https://www.termcat.cat/ca/diccionaris-en-linia/228/fitxa/MzU0MjAwNA==>. Darrera visita el 2 de setembre de 2023.

Discriminar. Real Acadèmia Espanyola. 2023. Disponible a: <https://dle.rae.es/discriminar>. Darrera visita el 2 de setembre de 2023.

DOBBE, R., DEAN, S., GILBERT, T. K., i KOHLI, N., “A Broader View on Bias in Automated Decision-Making: Reflecting on Epistemology and Dynamics”, 2018, pàg. 1-5. Disponible a: <https://arxiv.org/> Darrera visita el 2 de setembre de 2023.

EDRI *et al.*, “An EU Artificial Intelligence Act for Fundamental Rights. A civil Society Statement”, 30 November 2021. Disponible a: <https://edri.org/wp-content/uploads/2021/12/Political-statement-on-AI-Act.pdf> Darrera visita el 2 de setembre de 2023.

Equitat. Diccionari en línia. Enciclopèdia Catalana. 2023. Disponible a: <https://www.enciclopedia.cat/ec-gdlc-e00054409.xml>. Darrera visita el 2 de setembre de 2023.

ETICAS FOUNDATION, Observatory of Algorithms with Social Impact. Disponible a: <https://eticas.tech/es/oasi>. Darrera visita el 2 de setembre de 2023.

FUNDACIÓ TIC SALUT SOCIAL, “Guia sobre l’explicabilitat en la Intel·ligència Artificial”, 2022. Disponible a: <https://ticsalutsocial.cat/wp-content/uploads/2023/01/230226-Informe-Explicabilitat.pdf>. Darrera visita el 2 de setembre de 2023.

FERNANDEZ SAMANIEGO, J. i PIÑAR GUZMÁN, B., “Las acciones colectivas en el marco del RGPD: una perspectiva desde el Derecho Civil español”, Diario La Ley, núm. 26, Sección Ciberderecho, 11 de Febrero de 2019. Disponible a: <https://diariolaley.laleynext.es/Content/Documento.aspx?params=H4sIAAAAAAAAAEA MtMSbF1CTEAAiNLAwMLc7Wy1KLizPw8WyMDQ0sDI0NDkEBmWqVLfnJIZUG>

[qbVpiTnEqAJisSbM1AAAAWKE](#) Darrera visita el 2 de setembre de 2023.

Gènere. TERMCAT Centre de Terminologia. 2023. Disponible a: <https://www.termcat.cat/es/cercaterm/fitxa/MjcwMTQ1NA%3D%3D>. Darrera visita el 2 de setembre de 2023.

GONZÁLEZ CABAÑAS, J., CUEVAS, A. i CUEVAS, R., “Facebook Use of Sensitive Data for Advertising in Europe”, 2018, pàg. 1-15. Disponible a: <https://arxiv.org/>. Darrera visita el 2 de setembre de 2023.

GOVERNMENT OF CANADA, Algorithmic Impact Assessment Tool. Disponible a: <https://www.canada.ca/en/government/system/digital-government/digital-government-innovations/responsible-use-ai/algorithmic-impact-assessment.html>. Darrera visita el 2 de setembre de 2023.

GOVERNMENT OF THE NETHERLANDS, Fundamental Rights and Algorithms Impact Assessment (FRAIA). Disponible a: <https://www.government.nl/documents/reports/2021/07/31/impact-assessment-fundamental-rights-and-algorithms> Darrera visita el 2 de setembre de 2023.

HOLLAND, S., HOSNY, A., NEWMAN, S., JOSEPH, J. i CHMIELINSKI, K., “The Dataset Nutrition Label: A Framework To Drive Higher Data Quality Standards”, 2018, pàg. 1-21. Disponible a: <https://arxiv.org/>. Darrera visita el 2 de setembre de 2023.

Igualdad. Diccionario de la Real Academia Española en línea. 2023. Disponible a <https://dle.rae.es/igualdad>. Darrera visita el 2 de setembre de 2023.

Igualtat d'oportunitats. Diccionari en línia. Enciclopèdia Catalana. 2023. Disponible a: <https://www.enciclopedia.cat/ec-gec-0517880.xml>. Darrera visita el 2 de setembre de 2023.

Interés general. Diccionario panhispánico del español jurídico. 2023. Disponible a: <https://dpej.rae.es/lema/inter%C3%A9s-general>. Darrera visita el 2 de setembre de 2023.

Intersectionality. Cambridge Dictionary. Cambridge University Press. 2023. Disponible a: <https://dictionary.cambridge.org/es/diccionario/ingles/intersectionality>. Darrera visita el 2 de setembre de 2023.

MARTÍNEZ MARTÍNEZ, R., “Una aproximación crítica al Reglamento General de Protección de Datos”, Diario La Ley, núm. 10294, Sección Tribuna, 25 de Mayo de 2023.

Disponible

a:

https://diariolaley.laleynext.es/Content/Documento.aspx?params=H4sIAAAAAAAAAEAE1QTWvDMAz9NfMIMLJIFHrwJQ2UQRmly3ZXbS0xc6zMlrPk309JGZtBWNL74NlfGePS4sz6LUABY6TZDWDcXS7LjyoUJq4dWnZGYF9csPMwYGAqjhgwyspicY7EaP5UsmqAKam0BArLoNuYUTFck66eFBjO4BsyeuvdhC1cV4CixVgvulRMDP6CST887lXq6fsFJtcBOwo1xJuds1Y3bSmn2u135U5NGJM9LvrJCCq3nX9SYpv_IQQT-XGDvXJJQaxz57BUrqHNM7Kh0_J9Lqx_ivqzCyuVw43TBkvt7wPD-Ax2N88MI5-uZCX1Ns8UhK_vH7WczhApJzQ6_IH3dTnGnIBAAA=WKE Darrera visita el 2 de setembre de 2023.

MINISTERIO DE JUSTICIA DE ESPAÑA i MINISTERIO DE CONSUMO DE ESPAÑA, “Anteproyecto de Ley de acciones de representación para la protección de los intereses colectivos de los consumidores”, 2023. Disponible a:

<https://www.mjusticia.gob.es/es/AreaTematica/ActividadLegislativa/Documents/Anteproyecto%20de%20Ley%20acciones%20representativas.pdf> Darrera visita el 2 de setembre de 2023.

Multiple discrimination. European Institute for Gender Equality. 2023. Disponible a: <https://eige.europa.eu/thesaurus/terms/1297>. Darrera visita el 2 de setembre de 2023.

NANDO (New Approach Notified and Designated Organisations), Base de dades de la Unió Europea en línia. Disponible a: <https://ec.europa.eu/growth/tools-databases/nando/index.cfm>. Darrera visita el 2 de setembre de 2023.

NORTHCUTT, C.G. *et al.*, “Pervasive Label Errors in Test Sets Destabilize Machine Learning Benchmarks”, 2021, pàg. 1-24. Disponible a: <https://arxiv.org/> Darrera visita el 2 de setembre de 2023.

Opacitat. Diccionari en línia. Enciclopèdia Catalana. 2023. Disponible a: <https://www.diccionari.cat/GDLC/opacitat>. Darrera visita el 2 de setembre de 2023.

Principio de igualdad. Diccionario jurídico de derecho en línea. Enciclopedia Jurídica. 2023. Disponible a <http://www.encyclopedia-juridica.com/d/principio-de-igualdad/principio-de-igualdad.htm>. Darrera visita el 2 de setembre de 2023.

Sexe. Gran enciclopèdia catalana. 2023. Disponible a: https://www.enciclopedia.cat/cerca/gec?search_api_fulltext=sexe&field_faceta_cerca_1=All. Darrera visita el 2 de setembre de 2023.

U.S. FOOD AND DRUG ADMINISTRATION (FDA), “Artificial intelligence machine learning and (AI/ML) software as a medical Device (SaMD) action plan”, 2021. Disponible a: <https://www.fda.gov/medical-devices/software-medical-device-samd/artificial-intelligence-and-machine-learning-software-medical-device> Darrera visita el 2 de setembre de 2023.

WOODS, D., “LinkedIn’s Monica Rogati on “What Is a Data Scientist?””, *Forbes*, 27 de novembre de 2011. Disponible a: <https://www.forbes.com/sites/danwoods/2011/11/27/linkedins-monica-rogati-on-what-is-a-data-scientist/>. Darrera visita el 2 de setembre de 2023.

JURISPRUDÈNCIA I ALTRES RESOLUCIONS

1. Internacional

TEDH (Gran Sala), sentència de 6 d'abril de 2000, 34369/97, Thlimmenos contra Grècia.

TEDH (Gran Sala), sentència de 5 de juliol de 2005, 43577/98 i 43579/98, Nachova i altres contra Bulgària.

TEDH (Secció Quarta), sentència de 12 de desembre de 2006, 13378/05, Burden contra el Regne Unit.

TEDH (Gran Sala), sentència de 13 de novembre de 2007, 57325/00, D.H. i altres contra la República Txeca.

TEDH (Secció Quarta), sentència de 27 de novembre de 2007, 77782/01, Luczak contra Polònia.

TEDH (Gran Sala), sentència de 4 de desembre de 2008, 30562/04 i 30566/04, S. and Marper contra Regne Unit.

TEDH (Secció Tercera), sentència de 24 de juliol de 2012, 47159/08, B. S. contra Espanya.

TEDH (Secció Primera), sentència de 3 d'octubre de 2013, 552/10, I.B. contra Grècia.

TEDH (Secció Segona), sentència de 2 de desembre de 2014, 61960/08, Emel Boyraz contra Turquia.

TEDH (Secció Segona), sentència de 2 de febrer de 2016, 7186/09, Di Trizio contra Suïssa.

TEDH (Secció Segona), sentència de 22 de març de 2016, 23682/13, Guberina contra Croàcia.

TEDH (Gran Sala), sentència de 24 de maig de 2016, 38590/10, Biao contra Dinamarca.

TEDH (Secció Segona), sentència de 28 de març de 2017, 25536/14, Skorjanec contra Croàcia.

TEDH (Secció Quarta), sentència de 25 de juliol de 2017, 17484/15, Carvalho Pinto de Sousa Morais contra Portugal.

2. Unió Europea

TJCE, sentència de 8 d'abril de 1976, Assumpte 43-75, Gabrielle Defrenne contra Société anonyme belge de navigation aérienne Sabena.

TJCE, sentència de 31 de març de 1981, C-96/80, JP Jenkins contra Kingsgate (Clothing Productions) Ltd.

TJCE, sentència de 13 de maig de 1986, C-170/84, Bilka - Kaufhaus GmbH contra Karin Weber von Hartz.

TJCE, sentència de 17 d'octubre de 1989, C- 109/88, Handels-og Kontorfunktionærernes Forbund i Danmark contra Dansk Arbejdsgiverforening en nom de Danfoss.

TJCE, sentència de 8 de novembre de 1990, C-177/88, Elisabeth Johanna Pacifica Dekker contra Stichting Vormingscentrum voor Jong Volwassenen (VJV-Centrum) Plus.

TJCE, sentència de 27 d'octubre de 1993, C-127/92, Dr. Pamela Mary Enderby contra Frenchay Health Authority and Secretary of State for Health.

TJCE, sentència de 17 d'octubre de 1995, C-450/93, Eckhard Kalanke i Freie Hansestadt Bremen.

TJCE, sentència de 9 de febrer de 1999, C-167/97, Regina contra Secretary of State for Employment, ex parte Nicole Seymour-Smith y Laura Pérez.

TJCE (Sala Sisena), sentència d'11 de setembre de 2003, C-77/02, Erika Steinicke contra Bundesanstalt für Arbeit.

TJCE (Sala Cinquena), sentència de 23 d'octubre de 2003, C-4/02, Hilde Schönheit

contra Stadt Frankfurt am Main.

TJCE, sentència de 13 gener de 2004, C-256/01, Debra Allonby contra Accrington & Rossendale College, Education Lecturing Services, trading as Protocol Professional and Secretary of State for Education and Employment.

TJCE (Sala Primera), sentència de 27 de maig de 2004, C-285/02, Edeltraud Elsner-Lakeberg contra Land Nordrhein-Westfalen.

TJCE (Gran Sala), sentència de 8 de juny de 2004, C-220/02, Österreichischer Gewerkschaftsbund, Gewerkschaft der Privatangestellten contra Wirtschaftskammer Österreich.

TJCE, sentència de 10 de març de 2005, C-196/02, Vasiliki Nikoloudi contra Organismos Tilepikoinonion Ellados AE.

TJCE (Gran Sala), sentència de 22 de novembre de 2005, C-144/04, Weiner Mangold contra Rüdiger Helm.

TJCE (Gran Sala), sentència d'11 de juliol de 2006, C-13/05, Sonia Chacón Navas contra Eurest Colectividades SA.

TJCE (Gran Sala), sentència d'1 d'abril de 2008, C-267/06, Tadao Maruko contra Versorgungsanstalt der deutschen Bühnen.

TJCE, sentència de 10 de juliol de 2008, C-54/07, Centrum voor gelijkheid van kansen en voor racismebestrijding contra Firma Feryn NV.

TJCE (Gran Sala), sentència de 17 de juliol de 2008, C-303/06, S. Coleman contra Attridge Law and Steve Law.

TJUE (Sala Tercera), sentència de 5 de març de 2009, C-388/07, The Queen, on the application of The Incorporated Trustees of the National Council for Ageing (Age Concern England) contra Secretary of State for Business, Enterprise and Regulatory Reform.

TJUE (Gran Sala), sentència de 12 d'octubre de 2010, C-499/08, Ingeniørforeningen i Danmark, acting on behalf of Ole Andersen contra Region Syddanmark.

TJUE, sentència de 18 de novembre de 2010, C-356/09, Pensionsversicherungsanstalt contra Christine Kleist.

TJUE (Gran Sala), sentència d'1 de març de 2011, C-236/09, Association Belge des Consommateurs Test-Achats ASBL y otros contra Conseil des ministres.

TJUE (Sala Segona), sentència de 21 de juliol de 2011, C-104/10, Patrick Kelly contra National University of Ireland (University College, Dublin).

TJUE (Sala Segona), sentència de 19 d'abril de 2012, C-415/10, Galina Meister contra Speech Design Carrier Systems GmbH.

TJUE, sentència de 22 de novembre de 2012, C-385/11, Isabel Elbal Moreno e Instituto Nacional de la Seguridad Social (INSS), Tesorería General de la Seguridad Social (TGSS).

TJUE (Sala Tercera), sentència de 25 d'abril de 2013, C-81/12, Asociația Accept contra Consiliul Național pentru Combaterea Discriminării.

TJUE (Gran Sala), sentència de 18 de març de 2014, C-363/12, Z. contra A Government department and The Board of management of a community school.

TJUE (Gran Sala), sentència de 13 de maig de 2014, C-131/12, Google Spain S.L. contra Agència Espanyola de Protecció de dades.

TJUE (Sala Desena), sentència de 16 d'octubre de 2014, C-100/13, Comissió Europea contra la República Federal d'Alemanya.

TJUE (Gran Sala), sentència d'11 de novembre de 2014, C-530/13, Leopold Schmitzer contra Bundesministerin für Inneres.

TJUE (Sala Quarta), sentència de 18 de desembre de 2014, C-354/13, Fag og Arbejde (FOA) contra Kommunernes Landsforening (KL).

TJUE (Gran Sala), sentència de 16 de juliol de 2015, C-83/14, CHEZ Razpredelenie Bulgaria AD contra Komisia za zashtita ot diskriminatsia.

TJUE (Sala Tercera), sentència de 28 de juliol de 2016, C-191/15, Verein für Konsumenteninformation contra Amazon EU Sàrl.

TJUE (Sala Tercera), sentència de 27 d'octubre de 2016, C-613/14, James Elliott Construction Limited v Irish Asphalt Limited.

TJUE (Sala Primera), sentència de 24 de novembre de 2016, C443/15, David L. Parris contra Trinity College Dublin, Higher Education Authority, Department of Public Expenditure and Reform, Department of Education and Skills.

TJUE (Sala Segona), sentència de 9 de març de 2017, C-406/15, Petya Milkova contra Izpalnitelen direktor na Agentsiata za privatizatsia i sledprivatizatsionen kontrol, amb la intervenció de Varhovna administrativna prokuratura.

TJUE (Gran Sala), sentència de 14 de març de 2017, C-157/15, Samira Achbita i Centrum voor gelijkheid van kansen en voor racismebestrijding contra G4S Secure Solutions NV.

TJUE (Gran Sala), sentència de 14 de març de 2017, C-188/15, Asma Bougnaoui i Association de défense des droits de l'homme (ADDH) contra Micropole SA.

TJUE (Sala Primera), sentència de 6 d'abril de 2017, C-668/15, Jyske Finans A/S contra Ligebehandlingsnavnet.

TJUE (Sala Primera), sentència de 15 de novembre de 2018, C-457/17, Heiko Jonny Maniero contra Studienstiftung des deutschen Volkes Ev.

TJUE (Sala Tercera), sentència de 8 de maig de 2019, C-161/18, Violeta Villar Láiz contra Instituto Nacional de la Seguridad Social (INSS).

TJUE (Sala Tercera), sentència de 20 de juny de 2019, C-404/18, Jamina Hakelbracht y otros contra WTG Retail BVBA.

TJUE (Sala Tercera), sentència de 3 d'octubre de 2019, C- 274/18, Mino Schuch-Ghannadan contra Medizinische Universität Wien.

TJUE (Sala Primera), sentència de 12 de desembre de 2019, C-450/18, WA contra Instituto Nacional de la Seguridad Social (INSS).

TJUE (Gran Sala), sentència de 23 d'abril de 2020, C 507/18, NH contra Associazione Avvocatura per i diritti LGBTI - Rete Lenford.

TJUE (Sala Vuitena), sentència de 8 d'octubre de 2020, C-644/19, FT contra

Universitatea «Lucian Blaga» Sibiu, GS y otros, HS, Ministerul Educației Naționale.

TJUE (Gran Sala), sentència de 26 de gener de 2021, C-16/19, VL contra Szpital Kliniczny im. dra J. Babińskiego Samodzielny Publiczny Zakład Opieki Zdrowotnej w Krakowie.

TJUE (Gran Sala), sentència de 15 d'abril de 2021, C-30/19, Diskrimineringsombudsmannen contra Braathens Regional Aviation AB.

TJUE (Sala Tercera), sentència de 24 de febrer de 2022, C-389/20, CJ contra Tesorería General de la Seguridad Social (TGSS).

TJUE (Sala Tercera), sentència de 28 d'abril de 2022, C-319/20, Meta Platforms Ireland Limited, anteriorment Facebook Ireland Limited, contra Bundesverband der Verbraucherzentralen und Verbraucherverbände - Verbraucherzentrale Bundesverband e.V.

Conclusiones del abogado general Marco Darmon, presentadas el 15 de octubre de 1985, en el asunto C-170/84, Bilka - Kaufhaus GmbH contra Karin Weber von Hartz.

Conclusiones del abogado general Marco Darmon, presentadas el 14 de noviembre de 1989, en el asunto C-177/88, Elisabeth Johanna Pacifica Dekker contra Stichting Vormingscentrum voor Jong Volwassenen (VJV-Centrum) Plus.

Conclusiones del abogado general M. P. Léger, presentadas el 31 de mayo de 1995, en el asunto C-317/93, Inge Nolte contra Landesversicherungsanstalt Hannover.

Conclusiones del abogado general Dámaso Ruiz-Jarabo Colomer, presentadas el 6 de septiembre de 2007, en el asunto C-267/06, Tadao Maruko contra Versorgungsanstalt der deutschen Bühnen.

Conclusiones del abogado general Miguel Poiares Maduro, presentadas el 31 de enero de 2008, en el asunto C-303/06, S. Coleman contra Attridge Law i Steve Law.

Conclusiones del abogado general Miguel Poiares Maduro, presentadas el 12 de marzo de 2008, en el asunto, C-54/07, Centrum voor Gelijkheid van Kansen en voor Racismebestrijding contra Firma Feryn NV.

Conclusiones del abogado general Paolo Mengozzi, presentadas el 12 de enero de 2012, en el asunto C-415/10, Galina Meister contra Speech Design Carrier Systems GmbH.

Conclusiones de la abogada general Verica Trstenjak, presentadas el 28 de marzo de 2012, en el asunto C-171/11, Fra.bo SpA contra Deutsche Vereinigung des Gas- und Wasserfaches eV (DVGW) – Technisch- Wissenschaftlicher Verein.

Conclusiones de la abogada general Juliane Kokott, presentadas el 12 de marzo de 2015, en el asunto C-83/14, CHEZ Razpredelenie Bulgaria AD contra Komisia za zashtita ot diskriminatsia.

Conclusiones del abogado general M. Campos Sanchez-Bordona, presentadas el 28 de enero de 2016, en el asunto C-613/14, James Elliott Construction Limited contra Irish Asphalt Limited.

3. Espanya

Tribunal Constitucional (Pleno), sentencia 103/1983, de 22 de noviembre de 1983.

Tribunal Constitucional (Pleno), sentencia 75/1983, de 3 de agosto de 1983.

Tribunal Constitucional (Sala Segunda), sentencia 128/1987, de 16 de julio de 1987.

Tribunal Constitucional (sala primera), sentencia 47/1990, de 29 de marzo de 1990.

Tribunal Constitucional (Sala Primera), sentencia 41/1999, de 22 de marzo de 1999.

Tribunal Constitucional (Sala Segunda), sentencia 13/2001, de 29 de enero de 2001

Tribunal Constitucional (Sala Primera), sentencia 225/2001, de 26 de noviembre de 2001.

Tribunal Constitucional (Pleno), sentencia 253/2004, de 22 de diciembre de 2004.

Tribunal Constitucional (Sala Primera), sentencia 69/2007, de 16 de abril de 2007.

Tribunal Constitucional (Sala Segunda), sentencia 6/2011, de 14 de febrero de 2011.

Tribunal Constitucional (Pleno), sentencia núm. 76/2019, de 22 de mayo de 2019.

Tribunal Constitucional (Pleno), sentencia 91/2019, de 3 de julio de 2019.

Tribunal Constitucional (Sala Segunda), sentencia 1/2021, de 25 de enero de 2021.

Tribunal Supremo (Sala de lo Social, Cuarta), sentencia 3663/2000, de 4 de mayo del 2000 (recurso 3708/1999).

Tribunal Supremo (Sala de lo Social, Pleno), sentencia 805/2020, de 25 de septiembre de 2020 (recurso 4746/2019).

Audiència Provincial de Girona, auto de 16 de novembre de 2018 (recurs 890/2018).

Audiència Provincial de Girona, auto de 20 de desembre de 2018 (recurs 974/2018).

Juzgado Contencioso-Administrativo núm. 8, sentencia núm. 143/2021, de 30 de diciembre de 2021. Disponible a: <https://civio.app.box.com/s/ok4x3w0112iougk4cz6mt5ec73vdafb9> Darrera visita el 2 de setembre de 2023.

4. Altres països

ABRvS, 17 May 2017, ECLI:NL:RVS:2017:1259, Aerius I.

ABRvS, 18 July 2018, ECLI:NL:RVS:2018:2454, Aerius II.

(Bridges) contra CCSWP and SSHD [2019] EWHC 2341 (Admin). Disponible a: <https://www.judiciary.uk/wp-content/uploads/2020/08/R-Bridges-v-CC-South-Wales-ors-Judgment.pdf>. Darrera visita el 2 de setembre de 2023.

Conseil Constitutionnel, décision 2018-765DC, du 12 june 2018, considerant 65 a 72. Disponible a: <https://www.conseil-constitutionnel.fr/decision/2018/2018765DC.htm> Darrera visita el 2 de setembre de 2023.

Cour de cassation, décision 97-45258, du 28 mars 2000. Disponible a: https://www.courdecassation.fr/decision/6079b1a49ba5988459c52c35?search_api_fulltext=97-45258&op=Rechercher+sur+judilibre&date_du=2000-03-28&date_au=2000-03-28&judilibre_jurisdiction=cc&previousdecisionpage=&previousdecisionindex=&nextdecisionpage=&nextdecisionindex= Darrera visita el 2 de setembre de 2023.

State v. Loomis, 881, N.W.2d 749, 7532 (Wis. 2016).

Tribunal de Bolonya, sentència núm. 2949/2020, de 31 de desembre de 2020. Disponible a: <https://www.algoritmolegal.com/wp-content/uploads/2021/01/Sentencia-Bologna-Italia-Deliveroo-dic-2020-Original-italiano.pdf> Darrera visita el 2 de setembre de 2023.

Tribunal del Districte de La Haia, sentència de 5 de febrer de 2020, C/09/550982/ HA ZA 18-388, Case of SyRI. Disponible a: <https://uitspraken.rechtspraak.nl/inziendocument?id=ECLI:NL:RBDHA:2020:1878>. Darrera visita el 2 de setembre de 2023.

Tribunal Nacional Finlandès de No Discriminació i Igualtat. Resum de la decisió disponible a: <https://www.yvtltk.fi/en/index/opinionsanddecisions/decisions.html#> Darrera visita el 2 de setembre de 2023.

Tribunal Regional Administratiu de Lazio-Roma (Secció tercera bis), sentència núm. 3769, de 22 de març de 2017.

Tribunal Suprem dels Estats Units d'Amèrica del Nord, sentència de 8 de març de 1971, Griggs v. Duke Power Co., 401 U.S. 424. Disponible a: <https://supreme.justia.com/cases/federal/us/401/424/>. Darrera visita el 2 de setembre de 2023.

Tribunal Suprem dels Estats Units d'Amèrica del Nord, sentència de 24 de febrer de 1976, Mathews v. Eldridge, 424 U.S. 319, 333. Disponible a: <https://supreme.justia.com/cases/federal/us/424/319/>. Darrera visita el 2 de setembre de 2023.

5. Resolucions de cossos administratius

COMISSIÓ CATALANA DE GARANTIA DEL DRET D'ACCÉS A LA INFORMACIÓ PÚBLICA, Resolució de 21 de setembre de 2016, d'estimació de les Reclamacions 123/2016 i 124/2016 (acumulades).

COMMISSION D'ACCÈS AUX DOCUMENTS ADMINISTRATIFS, Avis 20144578, du 8 janvier 2015.

COMMISSION D'ACCÈS AUX DOCUMENTS ADMINISTRATIFS, Avis 20180276, du 19 avril 2018.

CONSEJO DE TRANSPARENCIA Y BUEN GOBIERNO, Resolución 701/2018, de 18 de febrero de 2019. Disponible a: <https://civio.app.box.com/s/0r9a58a10pdt2t08lnemlmc22sd7t31> Darrera visita el 2 de setembre de 2023.

LEGISLACIÓ

1. Internacional

Declaració de Drets Humans de les Nacions Unides

Pacte Internacional dels Drets Civils i Polítics

Conveni Europeu de Drets Humans

Convenio núm. 108 del Consejo de Europa para la protección de las personas con respecto al tratamiento automatizado de datos de carácter personal, hecho en Estrasburgo el 28 de enero de 1981.

Convenio núm. 108 del Consejo de Europa para la protección de las personas con respecto al tratamiento automatizado de datos de carácter personal, hecho en Estrasburgo el 28 de enero de 1981, modificado en la sesión núm. 128 del Comité de Ministros en Elnor el 18 de mayo de 2018.

Protocol 12 al Conveni Europeu de Drets Humans

2. Unió Europea

Tractat Constitutiu de la Comunitat Europea (TCE)

Tractat de la Unió Europea

Tractat de Funcionament de la Unió Europea

Carta de Drets Fonamentals de la Unió Europea

Reglamento (UE) 1025/2012 del Parlamento Europeo y del Consejo, de 25 de octubre de 2012, sobre la normalización europea, por el que se modifican las Directivas 89/686/CEE

y 93/15/CEE del Consejo y las Directivas 94/9/CE, 94/25/CE, 95/16/CE, 97/23/CE, 98/34/CE, 2004/22/CE, 2007/23/CE, 2009/23/CE y 2009/105/CE del Parlamento Europeo y del Consejo y por el que se deroga la Decisión 87/95/CEE del Consejo y la Decisión no 1673/2006/CE del Parlamento Europeo y del Consejo. DOUE, L 316/12, 14.11.2012.

Reglamento (UE) 2016/679 del Parlamento Europeo y del Consejo de 27 de abril de 2016 relativo a la protección de las personas físicas en lo que respecta al tratamiento de datos personales y a la libre circulación de estos datos y por el que se deroga la Directiva 95/46/CE (Reglamento general de protección de datos). DOUE, L 119/1, 4.5.2016.

Reglamento (UE) 2018/302 del Parlamento Europeo y del Consejo, de 28 de febrero de 2018, sobre medidas destinadas a impedir el bloqueo geográfico injustificado y otras formas de discriminación por razón de la nacionalidad, del lugar de residencia o del lugar de establecimiento de los clientes en el mercado interior y por el que se modifican los Reglamentos (CE) n.º 2006/2004 y (UE) 2017/2394 y la Directiva 2009/22/CE. DOUE, L 60/1, 2.3.2018.

Reglamento (UE) 2018/1725 del Parlamento Europeo y del Consejo, de 23 de octubre de 2018, relativo a la protección de las personas físicas en lo que respecta al tratamiento de datos personales por las instituciones, órganos y organismos de la Unión, y a la libre circulación de esos datos, y por el que se derogan el Reglamento (CE) n.º 45/2001 y la Decisión n.º 1247/2002/CE. DOUE, L 295/39, 21/11/2018.

Reglamento (UE) 2019/1020, del Parlamento Europeo y del Consejo, de 20 de junio de 2019, relativo a la vigilancia del mercado y la conformidad de los productos y por el que se modifican la Directiva 2004/42/CE y los Reglamentos (CE) núm. 765/2008 y (UE) núm. 305/2011. DOUE, L 169/1, 25.6.2019.

Reglamento (UE) 2022/868 del Parlamento Europeo y del Consejo de 30 de mayo de 2022 relativo a la gobernanza europea de datos y por el que se modifica el Reglamento (UE) 2018/1724 (Reglamento de Gobernanza de Datos). DOUE, L 152/1, 3.6.2022.

Reglamento (UE) 2022/1925 del Parlamento Europeo y del Consejo de 14 de septiembre de 2022 sobre mercados disputables y equitativos en el sector digital y por el que se modifican las Directivas (UE) 2019/1937 y (UE) 2020/1828 (Reglamento de Mercados

Digitales). DOUE, L 265/1, 12.10.2022.

Reglamento (UE) 2022/2065 del Parlamento Europeo y del Consejo de 19 de octubre de 2022 relativo a un mercado único de servicios digitales y por el que se modifica la Directiva 2000/31/CE (Reglamento de Servicios Digitales). DOUE, L 277/1, 27.10.2022.

Directiva 79/7/CEE del Consejo, de 19 de diciembre de 1978, relativa a la aplicación progresiva del principio de igualdad de trato entre hombres y mujeres en materia de seguridad social. DOCE, L 6/24, 10.01.1979.

Directiva 89/106/CEE del Consejo de 21 de diciembre de 1988 relativa a la aproximación de las disposiciones legales, reglamentarias y administrativas de los Estados Miembros sobre los productos de construcción. DOUE, L 040/12, 11.02.1989.

Directiva 92/85/CEE del Consejo, de 19 de octubre de 1992, relativa a la aplicación de medidas para promover la mejora de la seguridad y de la salud en el trabajo de la trabajadora embarazada, que haya dado a luz o en período de lactancia. DOCE, L 348/1, 28.11.1992.

Directiva 95/46/CE del Parlamento Europeo y del Consejo, de 24 de octubre de 1995, relativa a la protección de las personas físicas en lo que respecta al tratamiento de datos personales y a la libre circulación de estos datos. DOCE, L 281/31, 23.11.1995.

Directiva 97/80/CE del Consejo, de 15 de diciembre de 1997, relativa a la carga de la prueba en los casos de discriminación por razón de sexo. DOCE, L 14/6, 20.01.1998.

Directiva 97/81/CE del Consejo, de 15 de diciembre de 1997, relativa al Acuerdo marco sobre el trabajo a tiempo parcial concluido por la Unice, el Ceep y la Ces. DOCE, L 14/9, 20.01.1998.

Directiva 2000/43/CE del Consejo, de 29 de junio de 2000, relativa a la aplicación del principio de igualdad de trato de las personas independientemente de su origen racial o étnico. DOCE, L 180/22, 19.07.2000.

Directiva 2000/78/CE del Consejo, de 27 de noviembre de 2000, relativa al establecimiento de un marco general para la igualdad de trato en el empleo y la ocupación. DOCE, L 303/16, 2.12.2002.

Directiva del Consejo 2004/113/CE, de 13 de diciembre de 2004, por la que se aplica el principio de igualdad de trato entre hombres y mujeres al acceso a bienes y servicios y su suministro. DOUE, L 373/37, 21.12.2004.

Directiva 2006/54/CE del Parlamento Europeo y del Consejo, de 5 de julio de 2006, relativa a la aplicación del principio de igualdad de oportunidades e igualdad de trato entre hombres y mujeres en asuntos de empleo y ocupación (refundición), DOUE, L 204/23, 26.7.2006.

Directiva 2010/18/UE del Consejo, de 8 de marzo de 2010, por la que se aplica el Acuerdo marco revisado sobre el permiso parental, celebrado por BUSINESSEUROPE, la UEAPME, el CEEP y la CES. DOUE, L 68/13, 18.03.2010.

Directiva 2010/41/UE del Parlamento Europeo y del Consejo, de 7 de julio de 2010, sobre la aplicación del principio de igualdad de trato entre hombres y mujeres que ejercen una actividad autónoma. DOUE, L 180/1, 15.07.2010.

Directiva 2014/54/UE del Parlamento Europeo y del Consejo, de 16 de abril de 2014, sobre medidas para facilitar el ejercicio de los derechos conferidos a los trabajadores en el contexto de la libre circulación de los trabajadores. DOUE, L 128/8, 30.4.2014.

Directiva (UE) 2016/680 del Parlamento Europeo y del Consejo, de 27 de abril de 2016, relativa a la protección de las personas físicas en lo que respecta al tratamiento de datos personales por parte de las autoridades competentes para fines de prevención, investigación, detección o enjuiciamiento de infracciones penales o de ejecución de sanciones penales, y a la libre circulación de dichos datos y por la que se deroga la Decisión Marco 2008/977/JAI del Consejo. DOUE, L 119/89, 4.5.2016.

Directiva (UE) 2019/1158 del Parlamento Europeo y del Consejo, de 20 de junio de 2019, relativa a la conciliación de la vida familiar y la vida profesional de los progenitores y los cuidadores. DOUE, L 188/79, 12.07.2019.

Directiva (UE) 2020/1828 del Parlamento Europeo y del Consejo, de 25 de noviembre de 2020, relativa a las acciones de representación para la protección de los intereses colectivos de los consumidores, y por la que se deroga la Directiva 2009/22. DOUE, L 409/1, 4.12.2020.

Decisión Marco 2002/584/JAI del Consejo, de 13 de junio de 2002, relativa a la orden de detención europea y a los procedimientos de entrega entre Estados miembros. DOUE, L 190/1, 18.07.2002.

Decisión marco 2008/913/JAI del Consejo, de 28 de noviembre de 2008, relativa a la lucha contra determinadas formas y manifestaciones de racismo y xenofobia mediante el Derecho penal. DOUE, L 328/55, 6.12.2008.

Comunicación de la Comisión, «Guía azul» sobre la aplicación de la normativa europea relativa a los productos, de 2022. DOUE C 247/1, 29.6.2022.

Recomendación de la Comisión, de 11 de junio de 2013, sobre los principios comunes aplicables a los mecanismos de recurso colectivo de cesación o de indemnización en los Estados miembros en caso de violación de los derechos reconocidos por el Derecho de la Unión. DOUE, L 201/60, 26.7.2013.

Recomendación (UE) 2018/951 de la Comisión, de 22 de junio de 2018, sobre normas relativas a los organismos para la igualdad. DOUE, L 167/28, 4.7.2018.

2.1. Documents preparatius

COMISIÓN DE LAS COMUNIDADES EUROPEAS, *Propuesta de Directiva del Consejo relativa a la carga de la prueba en el ámbito de la igualdad de retribución y trato entre hombres y mujeres*, (COM (88) 269 final), Bruselas, 27 de mayo de 1988.

COMISIÓN DE LAS COMUNIDADES EUROPEAS, *Propuesta de Directiva del Consejo por la que se aplica el principio de igualdad de trato entre las personas independientemente de su religión o convicciones, discapacidad, edad u orientación sexual*, COM (2008) 426-final, Bruselas, 2 de julio de 2008.

COMISIÓN EUROPEA, *Propuesta de Reglamento del Parlamento europeo y del Consejo, por el que se establecen normas armonizadas en materia de inteligencia artificial (ley de inteligencia artificial) y se modifican determinados actos legislativos de la Unión*, COM(2021) 206, Bruselas, 24 de abril de 2021.

COMISIÓN EUROPEA, *Propuesta de Reglamento del Parlamento Europeo y del Consejo relativo a las máquinas y sus partes y accesorios*, COM(2021) 202 final, Bruselas, 21 de abril de 2021.

COMISIÓN EUROPEA, *Propuesta de Directiva del Parlamento europeo y del Consejo, relativa a la responsabilidad por productos defectuosos*, COM(2022) 495 final, Bruselas, 28 de septiembre de 2022.

COMISIÓN EUROPEA, *Propuesta de Directiva del Parlamento Europeo y del Consejo, relativa a la adaptación de las normas de responsabilidad no contractual a la inteligencia artificial*, COM (2022) 496 final, Bruselas, 28 de septiembre de 2022.

EUROPEAN COMMISSION, *Proposal for a Council Directive on standards for equality bodies in the field of equal treatment between persons irrespective of their racial or ethnic origin, equal treatment in the field of employment and occupation between persons irrespective of their religion or belief, disability, age or sexual orientation, equal treatment between women and men in matters of social security and in the access to and supply of goods and services, and deleting Article 13 of Directive 2000/43/EC and Article 12 of Directive 2004/113/EC*, COM(2022) 689 final, Brussels, 7 December 2022.

EUROPEAN COMMISSION, *Proposal for a Directive of the European Parliament and of the Council on standards for equality bodies in the field of equal treatment and equal opportunities between women and men in matters of employment and occupation, and deleting Article 20 of Directive 2006/54/EC and Article 11 of Directive 2010/41/EU*, COM(2022) 688 final, Brussels, 7 December 2022.

EUROPEAN PARLIAMENT, *Amendments adopted by the European Parliament on 14 June 2023 on the proposal for a regulation of the European Parliament and of the Council on laying down harmonised rules on artificial intelligence (Artificial Intelligence Act) and amending certain Union legislative acts (COM(2021)0206 – C9-0146/2021 – 2021/0106(COD))1*, P9_TA(2023)0236, Brussels, 14 June 2023.

3. Espanya

Constitución Española. BOE, núm. 311, de 29 de diciembre de 1978. España.

Ley Orgánica 3/2007, de 22 de marzo, para la igualdad efectiva de mujeres y hombres. BOE, núm. 71 de 23 de marzo de 2007. España.

Ley Orgánica 5/1985, de 19 de junio, del Régimen Electoral General. BOE, núm. 147, de 20 de junio 1985. España.

Ley Orgánica 6/1985, de 1 de julio, del Poder Judicial. BOE, núm. 157, de 02 de julio de 1985. España.

Ley Orgánica 3/2018, de 5 de diciembre, de Protección de Datos Personales y garantía de los derechos digitales. BOE, núm. 294, de 6 de diciembre de 2018. España.

Ley Orgánica 7/2021, de 26 de mayo, de protección de datos personales tratados para fines de prevención, detección, investigación y enjuiciamiento de infracciones penales y de ejecución de sanciones penales. BOE, núm. 126, de 27 de mayo de 2021. España.

Ley 29/1998, de 13 de julio, reguladora de la Jurisdicción Contencioso-administrativa. BOE, núm. 167, de 14 de julio de 1998. España.

Ley 1/2000, de 7 de enero, de Enjuiciamiento Civil. BOE, núm. 7, de 8 de enero del 2000. España.

Ley 62/2003, de 30 de diciembre, de medidas fiscales, administrativas y del orden social. BOE, núm. 313, de 31 de diciembre de 2003. España.

Ley 15/2007, de 3 de julio, de Defensa de la Competencia. BOE, núm. 159, de 4 de julio de 2007. España.

Ley 36/2011, de 10 de octubre, reguladora de la jurisdicción social. BOE, núm. 245, de 11 de octubre de 2011. España.

Ley 19/2013, de 9 de diciembre, de transparencia, acceso a la información pública y buen gobierno. BOE, núm. 295, 10 de diciembre de 2013. España.

Ley 39/2015, de 1 de octubre, del Procedimiento Administrativo Común de las Administraciones Públicas. BOE, núm. 236, de 2 d'octubre de 2015. España.

Ley 40/2015, de 1 de octubre, de Régimen Jurídico del Sector Público. BOE, núm. 236, de 2 de octubre de 2015. España.

Ley 7/2020, de 13 de noviembre, para la transformación digital del sistema financiero. BOE, núm. 300, de 14 de noviembre de 2020. España.

Ley 12/2021, de 28 de septiembre, por la que se modifica el texto refundido de la Ley del Estatuto de los Trabajadores, aprobado por el Real Decreto Legislativo 2/2015, de 23 de octubre, para garantizar los derechos laborales de las personas dedicadas al reparto en el ámbito de plataformas digitales. BOE, núm. 233, de 29 de septiembre de 2021. España.

Ley 22/2021, de 28 de diciembre, de Presupuestos Generales del Estado para el año 2022. BOE, núm. 312, de 29/12/2021. España.

Ley 15/2022, de 12 de julio, integral para la igualdad de trato y la no discriminación. BOE, núm. 167, de 13 de julio de 2022. España.

Real Decreto Legislativo 1/2007, de 16 de noviembre, por el que se aprueba el texto refundido de la Ley General para la Defensa de los Consumidores y Usuarios y otras leyes complementarias. BOE, núm. 287, de 30 de noviembre de 2007. España.

Real Decreto 1720/2007, de 21 de diciembre, por el que se aprueba el Reglamento de desarrollo de la Ley Orgánica 15/1999, de 13 de diciembre, de protección de datos de carácter personal. BOE, núm. 17, de 19 de enero de 2008. España.

Real Decreto Legislativo 1/2013, de 29 de noviembre, por el que se aprueba el Texto Refundido de la Ley General de derechos de las personas con discapacidad y de su inclusión social. BOE, núm. 289, de 03 de diciembre de 2013.

Real Decreto 568/2022, de 11 de julio, por el que se establece el marco general del banco de pruebas regulatorio para el fomento de la investigación y la innovación en el sector eléctrico. BOE, núm. 166, de 12 de julio de 2022. España.

Real Decreto 729/2023, de 22 de agosto, por el que se aprueba el Estatuto de la Agencia Española de Supervisión de Inteligencia Artificial. BOE, núm. 210, de 2 de septiembre de 2023. España.

Carta de Derechos Digitales

3.1. Documents preparatius

Enmiendas núm. 5 a 7, todas realizadas por el Grupo Parlamentario Confederal Unidos Podemos-En comú Podem-En Marea. Boletín Oficial de las Cortes Generales, Congreso de los Diputados, Serie A n.º 13-2, 18 de abril de 2018.

Proposición de Ley integral para la igualdad de trato y la no discriminación. BOCG, núm. 146-1, 29 de enero de 2021. España.

4. Comunitats Autònòmiques

Decreto-ley 2/2023, de 8 de marzo, de medidas urgentes de impulso a la inteligencia artificial en Extremadura. DOE, núm. 48, de 10 de marzo de 2023.

Llei 19/2020, de 30 de desembre, d'igualtat de tracte i no-discriminació. Diari Oficial de la Generalitat de Catalunya, núm. 8307, de 31 de desembre de 2020.

Llei 1/2022, de 13 d'abril, de la Generalitat, de Transparència i Bon Govern de la Comunitat Valenciana. DOGV, núm. 9323, de 22 d'abril de 2022.

5. Altres països

Décret n° 2017-330 du 14 mars 2017 relatif aux droits des personnes faisant l'objet de décisions individuelles prises sur le fondement d'un traitement algorithmique. JORF n°0064 du 16 mars 2017. France.

Décret n° 2017-888 du 6 mai 2017 relatif à l'action de groupe et à l'action en reconnaissance de droits prévues aux titres V et VI de la loi n° 2016-1547 du 18 novembre

2016 de modernisation de la justice du XXI^e siècle. JORF n°0109 du 10 mai 2017. France.

Equal Credit Opportunity Act, Pub. L. No. 93-495, 88 Stat. 1521 (1974) (codified as amended at 15 U.S.C. §§ 1691–1691f (2012)).

Fair Credit Reporting Act, Pub. L. No. 91-508, 84 Stat. 1127 (1970) (codified as amended at 15 U.S.C. §§ 1681–1681x (2012)).

Loi n° 78-17 du 6 janvier 1978 relative à l'informatique, aux fichiers et aux libertés. JORF n° 0006 du 7 janvier 1978. France.

Loi n° 2008-496 du 27 mai 2008 portant diverses dispositions d'adaptation au droit communautaire dans le domaine de la lutte contre les discriminations. JORF n°0123 du 28 mai 2008. France.

Loi n° 2016-1321 du 7 octobre 2016 pour une République numérique. JORF n°0235 du 8 octobre 2016. France.

Loi 2016-1547, du 18 novembre 2016, de modernisation de la justice du XXI^e siècle. JORF n° 0269 du 19 novembre 2016. France.

Loi du 30 juillet 2018 relative à la protection des personnes physiques à l'égard des traitements de données à caractère personnel. Moniteur Belge n° 209, 5 septembre 2018. Belgique.

Ordonnance n° 2015-1341 du 23 octobre 2015 relative aux dispositions législatives du code des relations entre le public et l'administration. JORF n°0248 du 25 octobre 2015. France.

Regulation B, 12 C.F.R. §§ 1002.1–.16 (2018).

Uniform Guidelines on Employment Selection Procedures adopted by the United States of North America Equal Employment Opportunity Commission, Civil Service Commission, Department of Labor, and Department of Justice , 1978.

Wet Structuur uitvoeringsorganisatie en inkomen, SUWI (Llei neerlandesa d'Organització d'Implementació i Estructura d'ingressos).

ANNEX. Taula esquemàtica de la combinació de la normativa europea antidiscriminació, la normativa europea de protecció de dades i la normativa europea respecte a la intel·ligència artificial

PROBLEMES	SOLUCIONS		
Capítol II i III	Capítol IV (Regulació antidiscriminació)	Capítol V (RGPD)	Capítol VI (PRIA)
<p>Llista exhaustiva d'atributs protegits:</p> <ul style="list-style-type: none"> ❖ Categories protegides no suficientment representatives. ❖ No inclusió de tipologies de diferenciació que reforcen desigualtats 	<ul style="list-style-type: none"> • Esquema de protecció més ampli. Llistes d'atributs obertes. <p>No poden incloure qualsevol forma de tractament injust.</p> <p><u>Inclusió només en cas que siguin:</u></p>		

PROBLEMES	SOLUCIONS		
Capítol II i III	Capítol IV (Regulació antidiscriminació)	Capítol V (RGPD)	Capítol VI (PRIA)
socials (p.e. estatus socioeconòmic).	<ul style="list-style-type: none"> - Atributs que es necessiti protegir perquè responen a una lògica present o futura de segregació (estatus socioeconòmic o material genètic). - Atributs que es relacionin de manera indissociable amb els atributs protegits comunament acceptats (color de 		

PROBLEMES	SOLUCIONS		
Capítol II i III	Capítol IV (Regulació antidiscriminació)	Capítol V (RGPD)	Capítol VI (PRIA)
	pell, lloc de procedència, etc.).		
<p>Protecció desigual segons l'atribut protegit en les Directives antidiscriminació europees:</p> <p>❖ S'adrecen a atributs protegits determinats, i no a totes les formes de discriminació.</p>	<ul style="list-style-type: none"> • Inclusió explícita de la protecció dels mitjans de comunicació i la publicitat en la Directiva 2000/43/CE • Equiparació de la protecció en relació amb l'àmbit material de totes les Directives (Proposta de Directiva Horitzontal de 2008). Que les proteccions 		

PROBLEMES	SOLUCIONS		
Capítol II i III	Capítol IV (Regulació antidiscriminació)	Capítol V (RGPD)	Capítol VI (PRIA)
<ul style="list-style-type: none"> ❖ La protecció no és la mateixa segons l'atribut protegit. ❖ S'apliquen majoritàriament en relacions laborals i no en altres àmbits. ❖ Proporcionen excepcions que suposen un gran buit en relació a les decisions preses mitjançant algoritmes 	<p>s'apliquin a l'àmbit laboral, l'accés a béns i serveis i l'estat de benestar.</p>		

PROBLEMES	SOLUCIONS		
Capítol II i III	Capítol IV (Regulació antidiscriminació)	Capítol V (RGPD)	Capítol VI (PRIA)
(com l'àmbit de la publicitat).			
<p>Irrellevància del caràcter grupal de de la discriminació (vegeu el capítol VII per una anàlisi de les possibilitats de l'acció col·lectiva en matèria de discriminació algorítmica).</p> <p>❖ La proposta de regulació “anticlassificació” (no</p>	<ul style="list-style-type: none"> • Adopció de la perspectiva antilibertatària • Reconeixement de la discriminació estructural, sensible als sistemes d'opressió. Per exemple, prohibició de l'ús de sistemes d'IA en contextos en els quals no es pugui assegurar que no s'estan perpetuant 		

PROBLEMES	SOLUCIONS		
Capítol II i III	Capítol IV (Regulació antidiscriminació)	Capítol V (RGPD)	Capítol VI (PRIA)
<p>utilitzar dades sensibles en el sistema algorítmic pot no funcionar si aquestes s'infereixen d'altres dades).</p> <p>❖ La discriminació causada per sistemes algorítmics: elaboració de perfils en base a sistemes d'opressió, que s'apliquen individualment.</p>	<p>sistemes d'opressió i subordinació.</p> <ul style="list-style-type: none"> • Reconeixement de la discriminació interseccional per facilitar la prova en casos de discriminació algorítmica al possibilitar l'aportació d'un únic índex de biaix interseccional en comptes d'índexs relatius a cada un dels atributs concurrents. Per exemple, 		

PROBLEMES	SOLUCIONS		
Capítol II i III	Capítol IV (Regulació antidiscriminació)	Capítol V (RGPD)	Capítol VI (PRIA)
<p>❖ Dret antidiscriminatori de la UE profundament marcat pel caràcter casuístic i individualitzat de l'aplicació judicial.</p>	<p>instauració d'una normativa que estableixi tants per cent de discriminació tolerable: si l'algoritme no és un 90 per cent precís per tots els grups de població, incloent grups interseccionals (p.e. dones, afroamericanes de més de 60 anys), no es podrà introduir en el mercat.</p>		

PROBLEMES	SOLUCIONS		
Capítol II i III	Capítol IV (Regulació antidiscriminació)	Capítol V (RGPD)	Capítol VI (PRIA)
	<p>També, instauració d'una normativa que estableixi que no es poden emprar sistemes en els quals les taxes d'error per tots els grups (incloent categories protegides combinades) superi l'1%.</p>		
<p>Dicotomia discriminació directa/discriminació indirecta:</p>	<ul style="list-style-type: none"> • Flexibilització del concepte de discriminació directa 		

PROBLEMES	SOLUCIONS		
Capítol II i III	Capítol IV (Regulació antidiscriminació)	Capítol V (RGPD)	Capítol VI (PRIA)
<ul style="list-style-type: none"> ❖ Atesa la dificultat d'accés i comprensió del funcionament del sistema algorítmic, dificultat de provar que la discriminació sigui directa. ❖ Atesa la manca d'elements probatoris, dificultat de provar discriminació indirecta. 	<p>per part de la jurisprudència del TJUE:</p> <ul style="list-style-type: none"> - El reconeixement de la discriminació per un factor indissociable a l'atribut protegit (Sentències TJUE: Dekker i Tadao Maruko) - La possibilitat d'establir la discriminació 		

PROBLEMES	SOLUCIONS		
Capítol II i III	Capítol IV (Regulació antidiscriminació)	Capítol V (RGPD)	Capítol VI (PRIA)
<p>❖ Probabilitat que la discriminació algorítmica recaigui en el concepte de discriminació indirecta (que permet la justificació objectiva).</p>	<p>directa sense referència comparativa (Sentència TJUE: Dekker)</p> <p>- El reconeixement de discriminació per associació (Sentències TJUE: Coleman, Chez i Hakelbracht i altres). Aquesta tipologia de discriminació també obre la possibilitat a</p>		

PROBLEMES	SOLUCIONS		
Capítol II i III	Capítol IV (Regulació antidiscriminació)	Capítol V (RGPD)	Capítol VI (PRIA)
	<p>admetre la discriminació per error en la discriminació algorítmica.</p> <ul style="list-style-type: none"> • En el cas de no poder-se provar discriminació directa, considerar la discriminació indirecta a través dels efectes de la pràctica (si els atributs utilitzats no es consideren indissociables a l'atribut protegit, atès 		

PROBLEMES	SOLUCIONS		
Capítol II i III	Capítol IV (Regulació antidiscriminació)	Capítol V (RGPD)	Capítol VI (PRIA)
	que llavors seria discriminació directa).		
<p>Responsabilitat per l'acció discriminatòria:</p> <ul style="list-style-type: none"> ❖ Es veu difuminada per la multiplicitat d'actors que intervenen en el disseny, el funcionament i l'ús del sistema algorítmic 	<ul style="list-style-type: none"> • Article 27 de la Llei 15/2022: responsabilitat de la persona física o jurídica que causi discriminació i les persones que contracten o prestadores de serveis quan la discriminació es produeixi en el seu àmbit d'organització i direcció i no hagin 	<ul style="list-style-type: none"> • Tal com es planteja la responsabilitat proactiva en l'RGPD: responsabilitat de la persona responsable de tractament si la discriminació es causada per una negligència seva (art. 24 RGPD). Si no és el cas, responsabilitat de l'organització (o persona) que emprà el sistema algorítmic 	<ul style="list-style-type: none"> • Tal com es planteja la responsabilitat en la PRIA: responsabilitat del proveïdor del sistema d'IA (article 46 PRIA i considerant 64). Es podrà evitar aquesta responsabilitat si s'han complert tots els requisits que s'estableixen en el Reglament, és a dir, si s'ha actuat diligentment. No es preveu

PROBLEMES	SOLUCIONS		
Capítol II i III	Capítol IV (Regulació antidiscriminació)	Capítol V (RGPD)	Capítol VI (PRIA)
	<p>aplicat mètodes i instruments per la detecció, aplicació de mesures preventives i l'articulació de mesures adequades pel cessament de situacions discriminatòries.</p> <ul style="list-style-type: none"> Article 44 de la Llei 19/2020: són responsables les persones físiques o jurídiques, públiques o privades, que, per acció 	<p>(a través del compromís assumit en codis de conducta, certificacions, marques i segells).</p> <p><u>Per tal que aquest sigui un règim de responsabilitat efectiu, es necessita:</u></p> <ul style="list-style-type: none"> Supervisió exhaustiva i periòdica per part d'agències públiques. Plantejament de la responsabilitat en base a la imprudència o negligència 	<p>responsabilitat per part de l'usuari del sistema d'IA.</p>

PROBLEMES	SOLUCIONS		
Capítol II i III	Capítol IV (Regulació antidiscriminació)	Capítol V (RGPD)	Capítol VI (PRIA)
	<p>o omissió, incorren en els supòsits d'infracció, i si una infracció és imputable a diverses persones i no és possible determinar el grau de participació de cadascuna en la comissió de la infracció, n'han de respondre solidàriament.</p>	<p>per adoptar decisions fundades en sistemes algorítmics amb resultats esbiaixats, donada la possibilitat de prevenir i rectificar biaixos.</p>	

PROBLEMES	SOLUCIONS		
Capítol II i III	Capítol IV (Regulació antidiscriminació)	Capítol V (RGPD)	Capítol VI (PRIA)
<p>Establiment d'un cas <i>prima facie</i></p> <ul style="list-style-type: none"> ❖ Desconeixement de la condició de persona discriminada, per tant, impossibilitat d'aportar indicis si no s'és conscient de la discriminació ❖ Alta probabilitat de discriminació múltiple, que causa l'atomització tant del 	<ul style="list-style-type: none"> • Elements útils TJUE. Cas “Feryn”, “Accept” i “Rete Lenford” - Possibilitat de presentar un cas de discriminació sense víctimes de discriminació expressament identificades - Legitimació d'entitats d'interès públic (si estan habilitades pel dret nacional) per 	<ul style="list-style-type: none"> • Article 80 RGPD: Representació per part d'entitats, organitzacions i associacions dels drets col·lectius de persones que s'han vist discriminades pel tractament de dades. <p>Més possibilitats de litigar i d'establir un cas <i>prima facie</i> de discriminació, si entitats, organitzacions i associacions també tenen la possibilitat de defensar drets col·lectius.</p>	

PROBLEMES	SOLUCIONS		
Capítol II i III	Capítol IV (Regulació antidiscriminació)	Capítol V (RGPD)	Capítol VI (PRIA)
<p>grup perjudicat com del grup comparador.</p> <p>❖ Dificultat, per consegüent, d'inversió de la càrrega de la prova. A més, qui decideix sobre la inversió són els tribunals nacionals (diversitat de casos)</p>	<p>exercir accions judicials a falta d'una persona demandant que al·legui haver patit discriminació.</p> <ul style="list-style-type: none"> - Suficiència d'un dany hipotètic • Doctrina de la discriminació <i>per se</i> dels sistemes algorítmics. <p><i>(Vegeu el capítol VII per una anàlisi de les possibilitats de</i></p>	<p><i>(Vegeu al capítol VII una anàlisi de les possibilitats de l'acció col·lectiva en matèria de protecció de dades per cessar l'ús de sistemes algorítmics que produeixin discriminació)</i></p>	

PROBLEMES	SOLUCIONS		
Capítol II i III	Capítol IV (Regulació antidiscriminació)	Capítol V (RGPD)	Capítol VI (PRIA)
	<i>millorar l'establiment d'un cas prima facie de discriminació a través de l'acció col·lectiva en matèria de discriminació algorítmica)</i>		
<p>Pràctica inexistència de <i>tertium comparationis</i></p> <p>❖ Es determina la legitimitat del grup comparador segons el context.</p>		<ul style="list-style-type: none"> • Ús de la informació de les avaluacions d'impacte algorítmiques (anant més enllà de les avaluacions de protecció de dades previstes en l'article 35 de l'RGPD) per poder comparar l'experiència individual amb 	<ul style="list-style-type: none"> • Ús de la documentació tècnica dels sistemes de risc alt per part d'un organisme nacional d'igualtat (que segons l'article 64.3 i el considerant 79 de la PRIA podrà sol·licitar la documentació tècnica que ha d'elaborar qualsevol proveïdor

PROBLEMES	SOLUCIONS		
Capítol II i III	Capítol IV (Regulació antidiscriminació)	Capítol V (RGPD)	Capítol VI (PRIA)
<p>❖ A causa de l'opacitat dels sistemes algorítmics, dificultat d'accés a informació que ajudi a establir un grup comparador.</p>		<p>l'impacte del sistema algorítmic en col·lectius oprimits determinats (p. e. dones, persones amb discapacitat, persones majors de 55 anys, etc.)</p>	<p>d'un sistema d'IA de risc alt) per valorar si existeix discriminació.</p> <p>Per tant, possible via per discutir la discriminació en els sistemes algorítmics:</p> <p>Queixa per part d'un particular a una autoritat nacional d'igualtat que pot tenir accés a la documentació tècnica d'un sistema d'IA específic (<i>qüestió estudiada amb profunditat en el capítol VIII</i>). A través d'aquesta</p>

PROBLEMES	SOLUCIONS		
Capítol II i III	Capítol IV (Regulació antidiscriminació)	Capítol V (RGPD)	Capítol VI (PRIA)
			<p>informació es podria dur a terme una acció col·lectiva de cessació de l'ús del sistema d'IA per part d'entitats legitimades que tenen més possibilitats d'accedir a informació sobre el sistema algorítmic que no pas les persones individualment considerades.</p> <ul style="list-style-type: none"> • En el cas de sistemes d'IA emprats en el sector públic, accés a la documentació tècnica en possessió de les autoritats supervisores competents per

PROBLEMES	SOLUCIONS		
Capítol II i III	Capítol IV (Regulació antidiscriminació)	Capítol V (RGPD)	Capítol VI (PRIA)
			part de les persones a través del dret d'accés a la informació pública (Llei 19/2023).
<p>Biaixos algorítmics:</p> <ul style="list-style-type: none"> ❖ L'RGPD no fa referència explícita als biaixos algorítmics en l'articulat ❖ La prohibició en l'article 9 de l'RGPD d'emprar dades en el tractament dades que 		<ul style="list-style-type: none"> • Qualitat de les dades: <ul style="list-style-type: none"> - Exigència d'exactitud (article 5 RGPD) - Control de les dades en l'RGPD: Dret d'accés (art. 15), rectificació (art. 16), supressió (art.17), limitació del tractament (art. 18), portabilitat (art.20), oposició (art. 21). 	<ul style="list-style-type: none"> • Qualitat de les dades: <ul style="list-style-type: none"> - Governança de dades (art. 10 PRIA) - Habilitació pel tractament de categories especials de dades per detectar biaixos en les dades (art. 10.5 PRIA) • Control en el disseny:

PROBLEMES	SOLUCIONS		
Capítol II i III	Capítol IV (Regulació antidiscriminació)	Capítol V (RGPD)	Capítol VI (PRIA)
<p>revelin informació sensible pot dificultar la detecció de biaixos algorítmics.</p>		<ul style="list-style-type: none"> - Considerant 71 RGPD. Implementar mesures tècniques sense especificar quines. <p>No és suficient. <u>Mesures legislatives necessàries:</u></p> <ul style="list-style-type: none"> - Especificar les mesures tècniques a aplicar - Considerar la inclusió dels biaixos algorítmics en l'articulat (però és evident la dificultat de regular els 	<ul style="list-style-type: none"> - Trasllet de mesures tècniques específiques en els estàndards harmonitzats per mitigar els biaixos algorítmics.

PROBLEMES	SOLUCIONS		
Capítol II i III	Capítol IV (Regulació antidiscriminació)	Capítol V (RGPD)	Capítol VI (PRIA)
		<p>biaixos implícits de les persones)</p> <p><u>Mesures no legislatives necessàries:</u></p> <ul style="list-style-type: none"> - Sensibilització social - Assegurar perspectives diverses i compromeses amb la no discriminació en els equips que dissenyen els sistemes algorítmics <p><i>Proposta raonable:</i></p> <p>Explicació per part de les persones que creen una base de dades de les</p>	

PROBLEMES	SOLUCIONS		
Capítol II i III	Capítol IV (Regulació antidiscriminació)	Capítol V (RGPD)	Capítol VI (PRIA)
		seves característiques i limitacions.	
<p>Opacitat, manca d'informació i transparència:</p> <ul style="list-style-type: none"> ❖ La informació que es dona en l'RGPD sempre és individual ❖ Opacitat inherent al funcionament del sistema algorítmic (a vegades inclús per a 	<ul style="list-style-type: none"> • Segons la sentència TJUE, “Meister”: - La denegació total d'accés a la informació constitueix un factor a tenir en compte en l'establiment d'un cas de discriminació <i>prima facie</i>. • Revaluar la proposta de directiva de la 	<ul style="list-style-type: none"> • Requisits d'accés i transparència que proposa l'RGPD. <u>Sota les condicions:</u> - Entendre “informació sobre la lògica del sistema” (articles 13.2 (f), 14.2 (g) i 15.1 (h) de l'RGPD, tal com es plantegen les explicacions contrafàctiques (és a dir, de manera senzilla i útil: informació sobre els factors 	<ul style="list-style-type: none"> • Informació relativa a la documentació tècnica (no accessible al públic però sí per a organismes d'igualtat nacionals) <p>Article 11 PRIA; Informació detallada en l'annex IV PRIA:</p> <p>2.b. Lògica general del sistema d'IA i dels algoritmes; les opcions clau de disseny (la justificació lògica i les hipòtesis plantejades, també respecte a les</p>

PROBLEMES	SOLUCIONS		
Capítol II i III	Capítol IV (Regulació antidiscriminació)	Capítol V (RGPD)	Capítol VI (PRIA)
<p>les persones que el dissenyen).</p> <ul style="list-style-type: none"> ❖ Drets de propietat intel·lectual i industrial que dificulten l'accés a informació. ❖ Dificultat de comprensió de la informació a la qual es pugui tenir accés per convertir-la en informació útil en un plet de discriminació. 	<p>Comissió (1988) en la qual s'establia que les parts en litigi podien disposar de qualsevol informació en poder de l'altra part necessària per fer valer els seus drets.</p>	<p>que han determinat la decisió)</p> <ul style="list-style-type: none"> - Ampliar l'abast d'aquestes decisions a totes les decisions automatitzades • Avaluacions d'impacte algorítmiques. <u>Sota les següents condicions</u>, aquestes poden ser efectives: <ul style="list-style-type: none"> - Publicitat - Participació de la societat civil 	<p>persones i grups de persones en relació amb les quals es pretén emprar el sistema).</p> <p>2.d. Metodologies i tècniques d'entrenament emprats.</p> <p>2.g. Informació sobre les dades de validació i prova emprades i les seves característiques principals, així com els efectes potencialment discriminatoris.</p> <p>3. Informació en relació als nivells de precisió per a les persones o grups de persones específics respecte als que preveu emprar el sistema (...) i</p>

PROBLEMES	SOLUCIONS		
Capítol II i III	Capítol IV (Regulació antidiscriminació)	Capítol V (RGPD)	Capítol VI (PRIA)
		<ul style="list-style-type: none"> • Transparència per identificar cada sistema algorítmic i possibilitat que les persones puguin agrupar-se segons quin sistema hagi tractat les seves dades 	<p>els resultats no desitjats previsibles i les fonts de risc per a la salut i la seguretat, els drets fonamentals i la discriminació en vista de la finalitat prevista del sistema d'IA.</p> <ul style="list-style-type: none"> • Base de dades pública (article 60 PRIA) - Constarà el nom, direcció i dades de contacte del proveïdor, el nom comercial del sistema d'IA i una descripció de la finalitat prevista per sistema.

PROBLEMES	SOLUCIONS		
Capítol II i III	Capítol IV (Regulació antidiscriminació)	Capítol V (RGPD)	Capítol VI (PRIA)
			<p>La persona que intueix haver estat discriminada podrà localitzar mitjançant quin sistema d'IA pot haver estat. D'aquesta manera, si fa una queixa a l'organisme d'igualtat nacional, aquest podrà sol·licitar la documentació tècnica del sistema específic que detalli la precisió dels sistema d'IA en grups de població concrets.</p>
Control poc exhaustiu del sistema algorítmic que pot		<ul style="list-style-type: none"> • Instruments que planteja l'RGPD: 	Sistemes d'IA de risc alt:

PROBLEMES	SOLUCIONS		
Capítol II i III	Capítol IV (Regulació antidiscriminació)	Capítol V (RGPD)	Capítol VI (PRIA)
<p>causar discriminació:</p> <ul style="list-style-type: none"> ❖ Avaluació d'impacte de protecció de dades insuficient per abordar la discriminació algorítmica. ❖ Sistema d'auditories plantejat en l'RGPD insuficient per detectar discriminació algorítmica. ❖ Possibilitat d'esdevenir processos merament 		<ul style="list-style-type: none"> - Auditories (art. 39 RGPD) - Avaluacions d'impacte algorítmiques (anar més enllà de les avaluacions d'impacte de protecció de dades de l'art. 35 RGPD) - Sistemes d'autoregulació que incloguin exigències en relació amb la discriminació (Codis de conducta -art. 40 RGPD-; Certificacions -art. 42 RGPD-; Marques -art. 42 RGPD-; i Segells -art. 42 RGPD-) <p><u>Sota les següents condicions:</u></p>	<ul style="list-style-type: none"> • Control <i>ex ante</i>: <ul style="list-style-type: none"> - Supervisió humana dels sistemes de risc alt (article 14 del PRIA) - Habilitació d'espais de prova (art. 53 PRIA) - Avaluació de conformitat del sistema d'IA (art. 43 PRIA) • Control <i>ex post</i>: <ul style="list-style-type: none"> - Sistema de seguiment posterior a la comercialització creat pel proveïdor (art. 61.1 PRIA)

PROBLEMES	SOLUCIONS		
Capítol II i III	Capítol IV (Regulació antidiscriminació)	Capítol V (RGPD)	Capítol VI (PRIA)
<p>formalistes i normatius sense més rellevància.</p> <ul style="list-style-type: none"> ❖ Control de les dades en l'RGPD en base a l'input però no en base a l'output. ❖ Sistema de control excessivament basat en la confiança en l'autoregulació de les entitats privades combinat amb autoritats públiques 		<ul style="list-style-type: none"> - Periodicitat - Que es duguin a terme de manera cíclica (anàlisi, monitorització i avaluació; anàlisi, monitorització, avaluació; i així successivament) - Publicitat 	<ul style="list-style-type: none"> - Supervisió dels sistemes d'IA de risc alt per part de les autoritats de vigilància de mercat (art. 59 PRIA)

PROBLEMES	SOLUCIONS		
Capítol II i III	Capítol IV (Regulació antidiscriminació)	Capítol V (RGPD)	Capítol VI (PRIA)
<p>amb poc recursos i poc poder.</p> <p><i>(Vegeu el capítol VIII del treball per a una anàlisi de les autoritats que poden controlar la discriminació algorítmica que produeixin els sistemes d'IA.)</i></p>			
<p>Perspectiva individual de la protecció</p>		<ul style="list-style-type: none"> • Ús de l'amenaça que proporciona que les persones puguin fer ús dels seus drets 	

PROBLEMES	SOLUCIONS		
Capítol II i III	Capítol IV (Regulació antidiscriminació)	Capítol V (RGPD)	Capítol VI (PRIA)
<p>❖ La perspectiva de protecció marcadament individual de l'RGPD dificulta l'abordatge de la discriminació algorítmica que presenta un element col·lectiu no eliminable (p. e. drets atorgats en l'article 22.3 de l'RGPD).</p> <p><i>(Vegeu el capítol VII del</i></p>		<p>individuals per motivar les organitzacions a participar en la governança col·laborativa.</p> <p><u>Riscs rellevants d'aquest règim per la protecció de drets fonamentals:</u></p> <ul style="list-style-type: none"> - Implicació del sector privat en la definició de conceptes clau de la regulació - Atesa la manca de transparència pública, que el sector públic no pugui 	

PROBLEMES	SOLUCIONS		
Capítol II i III	Capítol IV (Regulació antidiscriminació)	Capítol V (RGPD)	Capítol VI (PRIA)
<i>treball per una anàlisi de les possibilitats de l'acció col·lectiva en matèria de protecció de dades, incorporant l'element col·lectiu en la lluita envers la discriminació algorítmica).</i>		instaurar una rendició de comptes eficient	