

Deusto Journal of Human Rights

Revista Deusto de Derechos Humanos

No. 11/2023

DOI: <https://doi.org/10.18543/djhr122023>

ARTICLES / ARTÍCULOS

La discriminación algorítmica y su impacto en la dignidad de la persona y los derechos humanos. Especial referencia a los inmigrantes

Algorithmic discrimination and its impact on human dignity and human rights.
Special reference to immigrants

José Miguel Iturmendi Rubia

<https://doi.org/10.18543/djhr.2910>

Fecha de publicación en línea: diciembre de 2023

Copyright (©)

Deusto Journal of Human Rights / Revista Deusto de Derechos Humanos is an Open Access journal; which means that it is free for full and immediate access, reading, search, download, distribution, and reuse in any medium only for non-commercial purposes and in accordance with any applicable copyright legislation, without prior permission from the copyright holder (University of Deusto) or the author; provided the original work and publication source are properly cited (Issue number, year, pages and DOI if applicable) and any changes to the original are clearly indicated. Any other use of its content in any medium or format, now known or developed in the future, requires prior written permission of the copyright holder.

Derechos de autoría (©)

Deusto Journal of Human Rights / Revista Deusto de Derechos Humanos es una revista de Acceso Abierto; lo que significa que es de libre acceso en su integridad inmediatamente después de la publicación de cada número. Se permite su lectura, la búsqueda, descarga, distribución y reutilización en cualquier tipo de soporte sólo para fines no comerciales y según lo previsto por la ley; sin la previa autorización de la Editorial (Universidad de Deusto) o la persona autora, siempre que la obra original sea debidamente citada (número, año, páginas y DOI si procede) y cualquier cambio en el original esté claramente indicado. Cualquier otro uso de su contenido en cualquier medio o formato, ahora conocido o desarrollado en el futuro, requiere el permiso previo por escrito de la persona titular de los derechos de autoría.

Deusto Journal of Human Rights

ISSN: 2530-4275 • ISSN-e: 2603-6002, No. 12/2023, Bilbao

© Universidad de Deusto • <http://djhr.revistas.deusto.es/>

La discriminación algorítmica y su impacto en la dignidad de la persona y los derechos humanos. Especial referencia a los inmigrantes

Algorithmic discrimination and its impact on human dignity and human rights. Special reference to immigrants

José Miguel Iturmendi Rubia 

CUNEF Universidad. España

jmiturmendi@cunef.edu

<https://doi.org/10.18543/djhr.2910>

Fecha de recepción: 31.05.2023

Fecha de aceptación: 19.10.2023

Fecha de publicación en línea: diciembre de 2023

Cómo citar / Citation: Iturmendi, José M. 2023. «La discriminación algorítmica y su impacto en la dignidad de la persona y los derechos humanos. Especial referencia a los inmigrantes.» *Deusto Journal of Human Rights*, No. 12: 257-284. <https://doi.org/10.18543/djhr.2910>

Sumario: 1. Contexto del problema. 2. Objetivos de la investigación. 3. Discriminación algorítmica y derechos humanos. 3.1. Definición de la discriminación algorítmica. 3.2. Impacto en los derechos humanos. 3.3. Ejemplos de discriminación algorítmica. 4. Factores que influyen en la discriminación algorítmica. 4.1. Sesgos en los datos y modelos utilizados en los algoritmos. 4.2. Falta de transparencia y explicabilidad de los algoritmos. 4.3. Falta de diversidad en el desarrollo de algoritmos. 4.4. Falta de transparencia y responsabilidad en el diseño y uso de los algoritmos. 5. Regulación de la discriminación algorítmica. 5.1. Marco legal actual. 5.2. Necesidad de una regulación adecuada. 5.3. Principios de una regulación justa y equitativa. 6. Soluciones técnicas a la discriminación algorítmica. 6.1. Técnicas de inteligencia artificial para la eliminación de sesgos. 6.2. Evaluación de la calidad y equidad de los algoritmos. 7. Caso de estudio: Discriminación de los inmigrantes. 7.1. Impacto de la discriminación algorítmica en los inmigrantes. 7.2. Ejemplos de discriminación algorítmica en el contexto migratorio. 8. Reflexiones finales. 9. Recomendaciones para la regulación y uso justo de los algoritmos. Referencias bibliográficas.

Resumen: La discriminación algorítmica, también conocida como sesgo algorítmico, se refiere al fenómeno en el cual los algoritmos de inteligencia

artificial y aprendizaje automático perpetúan o incluso amplifican los prejuicios y sesgos humanos existentes en los datos con los que son entrenados. Este problema plantea serias preocupaciones en relación con los derechos humanos, especialmente en el contexto de los inmigrantes. Puede conducir a la exclusión social, la marginalización y la perpetuación de estereotipos negativos. Estos algoritmos opacos pueden dificultar la rendición de cuentas y la apelación de decisiones injustas. Es fundamental abordar este problema mediante la implementación de prácticas de desarrollo de algoritmos éticos, como la transparencia en el diseño y la recopilación de datos, la auditoría de algoritmos y la diversidad en los equipos de desarrollo. Además, se requiere una regulación efectiva que garantice la protección de los derechos humanos en el contexto de la discriminación algorítmica, especialmente para los inmigrantes.

Palabras clave: discriminación algorítmica, sesgo algorítmico, inteligencia artificial, aprendizaje automático, derechos humanos, inmigrantes.

Abstract: Algorithmic discrimination, also known as algorithmic bias, refers to the phenomenon in which artificial intelligence algorithms and machine learning perpetuate or even amplify existing human prejudices and biases in the data they are trained on. This issue raises serious concerns regarding human rights, especially in the context of immigrants. It can lead to social exclusion, marginalization, and the perpetuation of negative stereotypes. These opaque algorithms can hinder accountability and the appeal of unfair decisions. Addressing this problem requires the implementation of ethical algorithm development practices, such as transparency in design and data collection, algorithm auditing, and diversity in development teams. Additionally, effective regulation is needed to ensure the protection of human rights in the context of algorithmic discrimination, particularly for immigrants.

Keywords: Algorithmic discrimination, algorithmic bias, artificial intelligence, machine learning, human rights, immigrants.

1. Contexto del problema

La discriminación algorítmica constituye una de las más destacadas amenazas al principio de la dignidad de la persona en un ámbito caracterizado por un vertiginoso progreso científico-tecnológico en el que resulta necesario destacar el crecimiento y el desarrollo de las tecnologías de inteligencia artificial y aprendizaje automático en una amplia variedad de áreas. Sistemas de IA cuyas evidentes virtudes —mejoras de la eficiencia y de la precisión de muchos procesos, transformación del mundo, de la persona, de sus maneras de vivir...— no pueden llevarnos a abrazar acríticamente su implementación y uso ni negar no ya sus implicaciones últimas, sino incluso las más inmediatas, pues pueden perpetuar y ampliar la discriminación de determinados colectivos y los prejuicios existentes en determinados ámbitos y sectores de la vida social. Estos sesgos pueden tener graves consecuencias para un buen número de derechos humanos, como el derecho a la privacidad, la interdicción de la discriminación o la igualdad ante la ley (Asís 2020, 225-228).

Tal y como ha puesto de manifiesto Silvia Barona (2021a, 119), “los sistemas de IA son un verdadero peligro para la humanidad, porque nos recomiendan implícitamente los sistemas de valores de sus creadores humanos, sin hacer transparentes estas recomendaciones”.

En este contexto, se hace necesario analizar las dimensiones éticas y legales de la discriminación algorítmica y su impacto en los derechos humanos. Este artículo aborda este problema, centrándose en la discriminación de los inmigrantes. A través de la revisión crítica de la dispar y cada vez más ingente literatura, se analizan las principales formas de tratamiento injusto como consecuencia de la toma de decisiones algorítmicas automatizadas y las formas en que estos algoritmos enquistan la discriminación de diversos colectivos. Asimismo, nos ocupamos de detallar el impacto de los diversos marcos legales y éticos en la discriminación algorítmica y ofrecemos propuestas que sirvan a la protección de los derechos fundamentales.

La discriminación algorítmica es un problema serio y urgente que puede tener consecuencias graves para los derechos humanos, especialmente para los inmigrantes. Es importante abordar este problema de manera efectiva y desarrollar soluciones prácticas y sostenibles que promuevan la igualdad y la no discriminación en todos los ámbitos de la sociedad. Resulta pertinente traer a colación, a modo de ejemplo, algunos de los algoritmos que más han escandalizado a la sociedad y que deberían señalarnos el camino para prevenirlos, como el empleado en el área de salud de Estados Unidos, que claramente

discriminaba en los programas de atención sanitaria a las personas de raza negra (Salas 2019) o la estigmatización de esos afroamericanos con herramientas que predicen supuestos de maltrato infantil.

2. Objetivos de la investigación

El objetivo de esta investigación es analizar la relación entre la discriminación algorítmica y los derechos humanos, particularmente en lo que respecta al derecho a la no discriminación, el derecho a la privacidad y el principio de igualdad ante la ley. La discriminación algorítmica se refiere al uso de algoritmos de aprendizaje automático para tomar decisiones que afectan a las personas en áreas como el empleo, la vivienda, la educación y la atención médica. A medida que la inteligencia artificial se vuelve cada vez más ubicua en la sociedad, la discriminación algorítmica se ha convertido en un problema cada vez más relevante.

La investigación examinará los desafíos y perspectivas de la discriminación algorítmica en la era de la inteligencia artificial, así como las posibles soluciones y enfoques para abordar este problema (Blanco 2022, 245). Además, se analizará el impacto de la discriminación algorítmica en los grupos vulnerables, como los inmigrantes, y se identificarán las implicaciones éticas y legales de la discriminación algorítmica en relación con los derechos humanos.

El propósito de esta investigación es contribuir a la comprensión del problema de la discriminación algorítmica y proporcionar recomendaciones prácticas para abordar este asunto en el marco de los derechos humanos. Con esto en mente, se espera que esta investigación sea de interés para todos aquellos que trabajan en el ámbito de las investigaciones jurídicas, incluyendo investigadores, responsables políticos, defensores de los derechos humanos y cualquier persona interesada en la intersección entre la inteligencia artificial y los derechos humanos.

3. Discriminación algorítmica y derechos humanos

3.1. *Definición de la discriminación algorítmica*

La discriminación algorítmica es un fenómeno que se produce cuando los sistemas de inteligencia artificial utilizan algoritmos y técnicas de aprendizaje automático para tomar decisiones que afectan

a las personas, pero que resultan discriminatorias. Estos sistemas pueden utilizar diferentes fuentes de datos para llevar a cabo sus decisiones, como datos históricos, patrones de comportamiento o datos de encuestas, entre otros. Sin embargo, estos datos pueden contener sesgos que se traducen en discriminación, lo que se traduce en la exclusión de ciertos grupos de personas, como minorías étnicas, inmigrantes, mujeres o personas con discapacidad, entre otros (Danesi 2022, 256-258). Es importante señalar que esta forma de discriminación no es intencional, sino que es el resultado de la falta de diversidad en los datos utilizados para entrenar los sistemas de inteligencia artificial.

Convenimos con la profesora Barona (2021a, 131) que “una de las formas para evitarlos es incrementar el volumen y la variedad de la información a partir de la cual los algoritmos son capaces de ‘aprender’ o de entrenarse favoreciendo de este modo la eliminación de posibles sesgos. O lo que es lo mismo, habrá que mantener la tensión constante entre privacidad y acceso a la información para el desarrollo y perfeccionamiento de estos sistemas. Se pretende incidir en lo que se denomina la ‘explicabilidad o la trazabilidad’ de la IA, que es la posibilidad de que el proceso de toma de decisiones de un algoritmo sea accesible, tratando con ello de evitar las denominadas ‘cajas negras’”.

En este contexto, el objetivo de esta investigación es analizar el fenómeno de la discriminación algorítmica en el contexto de la toma de decisiones sobre los derechos humanos de los inmigrantes. En particular, se busca comprender cómo se produce la discriminación algorítmica en la toma de decisiones sobre inmigración, qué consecuencias tiene para los derechos humanos de los inmigrantes, y qué medidas se pueden tomar para prevenir y abordar este problema. Para ello, se realizará una revisión sistemática de la literatura existente sobre el tema, así como un análisis crítico de los principales estudios empíricos que han investigado la discriminación algorítmica en el contexto de la inmigración, para tratar de evitar los sesgos con perfiles ideológicos, sociológicos, racistas o sexistas.

3.2. *Impacto en los derechos humanos*

La discriminación algorítmica puede tener un impacto significativo en los derechos humanos, especialmente en lo que respecta al respeto a la dignidad humana, el principio de igualdad y la interdicción de la discriminación. Los algoritmos que se utilizan para tomar decisiones

pueden estar sesgados y basados en datos históricos que reflejan la discriminación sistemática de ciertos colectivos, como las minorías étnicas o los inmigrantes. Esto puede llevar a decisiones discriminatorias y a la exclusión de ciertos grupos al acceso de oportunidades de diversa índole y de variados recursos. Además, la opacidad de los algoritmos puede dificultar la responsabilidad y la transparencia en la toma de decisiones, lo que puede obstaculizar la capacidad de los individuos para impugnar decisiones injustas o discriminatorias. Por lo tanto, es crucial abordar la discriminación algorítmica como un problema de derechos humanos y trabajar hacia soluciones que promuevan la igualdad y la no discriminación.

El Informe de Van Est y Gerritsen (2017), publicado por la Fundación Rathenau y adoptado por la Asamblea Parlamentaria del Consejo de Europa (PACE), pone el foco en este punto, al señalar la posible erosión que la inteligencia artificial puede producir en los derechos humanos, concluyendo en su punto octavo que se debe aclarar la importancia de garantizar la responsabilidad de quienes programan algoritmos con el fin de que se asuma las consecuencias de las decisiones automatizadas, evitando que supongan una injerencia en los derechos humanos y no generen discriminación a través de los sesgos.

3.3. Ejemplos de discriminación algorítmica

La discriminación algorítmica puede manifestarse en diversos contextos, y existen varios ejemplos documentados de su impacto. Por ejemplo, en el ámbito laboral, se han reportado casos en los que los algoritmos utilizados para seleccionar candidatos a un empleo han provocado la exclusión de grupos minoritarios, como personas de ciertas razas o géneros (Corrales 2022, 418). Como ha sido puesto de manifiesto por Josu Andoni Eguíluz "puede establecerse la discriminación directa incluso si no existe ningún reclamante identificable que alegue haber sido víctima de dicha discriminación. En el asunto Feryn, el Tribunal de Justicia de la Unión Europea (TJUE) concluyó que el hecho de que un empleador declare públicamente que no contratará a trabajadores de determinado origen étnico o racial constituye una discriminación directa en la contratación, en el sentido de la Directiva (2000/43/CE), ya que declaraciones de esa clase pueden disuadir firmemente a determinados candidatos de la solicitud del empleo y por tanto dificultar su acceso al mercado de trabajo" (Eguíluz 2020, 334).

Asimismo, en el ámbito financiero, se han denunciado prácticas de discriminación algorítmica en la concesión de créditos, donde los algoritmos utilizados para evaluar la solvencia de un solicitante pueden tener en cuenta factores irrelevantes o discriminatorios, como la raza o el género. Destaca Eguíluz que “hay algoritmos que clasifican a los ciudadanos en diferentes categorías de solvencia, en función de las probabilidades de que devuelva un préstamo, lo cual en un primer momento no es algo ilegal. No obstante, el cálculo de esos modelos parte de su historial crediticio, nivel de ingresos y otros datos que tienen un carácter especialmente protegido por el artículo nueve del RGPD, como son su raza o género. La puntuación conseguida por cada individuo determinará el tipo de interés que le ofrezcan los bancos, lo que en el caso de una hipoteca puede significar miles de euros anuales, algo que afecta sin duda al ámbito económico individual de cada ciudadano” (Eguíluz 2020, 357).

Otro ejemplo de discriminación algorítmica se ha documentado en el ámbito de la jurisdicción penal, donde los algoritmos utilizados para predecir la reincidencia de un delincuente pueden tener en cuenta factores que están correlacionados con la raza o el estatus socioeconómico, lo que puede conducir a la adopción de decisiones sesgadas y discriminatorias. Como destaca Silvia Barona, en relación con la justicia digital, “sus bondades pueden generar paz y bienestar social; sus maldades pueden propulsar discriminación, desigualdad, pervisión del sistema y aniquilación de las libertades y garantías. Solo desde la capacidad del ser humano y de la mirada jurídica seremos capaces de conseguir la humanización de los algoritmos y no la algoritmización de los humanos” (Barona 2021b, 59).

De igual manera, el profesor Pere Simón Castellano se ha referido a los sistemas expertos de inteligencia artificial judicial de valoración de riesgos (IAJVR) utilizados para, por ejemplo, detectar la reincidencia de un condenado. En este sentido, advierte que “la proyección de sesgos tan humanos como los basados en raza, género, religión o pobreza, sobre el algoritmo de los sistemas de IAJVR, debe evitarse a toda costa, puesto que puede tener un efecto multiplicador, propio de lo digital” (Simón 2021, 207).

Existen muchos países donde se desarrollan “programas predictivos de delitos mediante algoritmos capaces de procesar datos históricos individuales para diseñar mapas de delincuencia en distintas ciudades. En algunos de estos modelos, para calcular los índices, se pregunta por los antecedentes penales de amigos y familiares o por la primera vez que se tuvo trato con la policía y las respuestas, como es de esperar, varían sobre la base de circunstancias sociales, es decir, del nivel de

vida o de marginación social de unos barrios frente a otros. El LSI-R (*Level of Service Inventory-Revised*), por ejemplo, es un modelo matemático de peligrosidad y reincidencia basado en cuestionarios de preguntas. Si bien estas representaciones no pueden contener referencias a criterios como la raza porque serían ilegales, pueden incurrir igualmente en retroalimentaciones sesgadas. Un sujeto procedente de un barrio de 'alto riesgo' de reincidencia, donde es más probable que algún conocido haya tenido roces con la ley, se verá expuesto a condenas más largas y si pasa más tiempo en la cárcel, más difícil se tornará su reinserción en la vida laboral tras la salida. Con ello, también aumentarán las probabilidades objetivas de que vuelva a delinquir. Si nos atenemos a un estricto criterio matemático, resulta que el programa supo prever el resultado, pero, de hecho, el propio modelo utilizado habrá desempeñado un papel activo en causarlo" (Campione 2021, 132-133).

En el proceso penal es de vital importancia investigar sobre los algoritmos aparentemente neutrales que permiten el funcionamiento de la inteligencia artificial, dado que la persecución penal es por naturaleza sesgada, corriéndose el riesgo de automatizar dichos sesgos y producir sentencias masificadas regidas por probabilidades. La profesora López Pinilla, en un estudio sobre un algoritmo implementado en Colombia, señala que "el Fiscal Watson emplea una pequeña porción de la información de los delitos reportados, para arrojar una probabilidad de que alguien en concreto pueda ser responsable penalmente. Al ser una herramienta patentada por una compañía privada, una persona tendrá muy pocas probabilidades de controvertir el porcentaje arrojado por el algoritmo" (López Pinilla 2021, 147).

Son muchos los ejemplos que podemos poner del uso de estos algoritmos predictivos de reincidencia en el sistema penal. Así, HART (*Harm Assessment Risk Tool*), que predice el nivel de riesgo de los individuos de cometer delitos en un lapso de dos años, PSA (*Public Safety Assessment*), usado para determinar si el procesado puede permanecer en libertad hasta la etapa del juicio, teniendo en cuenta el nivel de riesgo de reincidencia o de no comparecencia, LSI-R (*Level of Service Inventory-Revised*), el cual se usa en los permisos de salida y la libertad condicional de los procesados, CAS (*Crime Anticipation System*), para predecir los delitos de alto impacto, NDAS (*National Data Analytics Solution*), para delitos violentos, o el más conocido COMPAS (*Correctional Offender Management Profiling for Alternative Sanctions*) usado para analizar el riesgo de reincidencia y las necesidades criminológicas del sujeto (Roa y Sanabria-Moyano 2022, 279-282).

De manera similar, en el ámbito de la atención médica, se han registrado casos en los que los algoritmos utilizados para determinar el tratamiento adecuado para un paciente han resultado en la exclusión o el tratamiento inadecuado de grupos minoritarios. Asimismo, la investigación con fármacos se realiza de forma frecuente con grupos etarios que no tienen nada que ver con la realidad (mayores, mujeres, diferentes etnias...). En otros casos es el propio algoritmo el que, basándose en algún criterio externo como la marca que hace un oncólogo en una mamografía sospechosa, establece un cribado que da como resultado la aparición de falsos positivos o, lo que es aún más grave, de falsos negativos (Febles 2012, 22).

Estos ejemplos muestran que la discriminación algorítmica puede tener consecuencias graves en diferentes áreas de la sociedad, y resaltan la importancia de abordar este problema de manera efectiva para promover la igualdad y la no discriminación en todos los ámbitos.

4. Factores que influyen en la discriminación algorítmica

4.1. *Sesgos en los datos y modelos utilizados en los algoritmos*

La discriminación algorítmica puede surgir como resultado de sesgos en los datos utilizados para entrenar y probar los algoritmos. Si partimos de datos históricos que contienen prejuicios y discriminación, los algoritmos pueden aprender y perpetuar estos prejuicios. Esto puede llevar a que los resultados de los algoritmos estén sesgados y que las decisiones basadas en ellos sean discriminatorias.

Los modelos utilizados en los algoritmos también pueden ser sesgados si están diseñados de manera que no tengan en cuenta ciertos factores, como la raza o la etnia. Si se utiliza un modelo que no tiene en cuenta estos factores, es más probable que se produzcan resultados discriminatorios.

Hay situaciones en las que se lleva a cabo una sobreexposición de imágenes de inmigrantes, con lo que esto conlleva para la vulneración de sus derechos de intimidad y protección de datos. Incluso este hecho es más obvio si pensamos en los menores que sirven como modelos para campañas de recogida de ayudas a través de ONGs. También se advierten sesgos en algoritmos asociados a cámaras de videovigilancia que devuelven mayores resultados de falsos positivos proporcionalmente al grado de oscuridad de su piel. Y, llevado al extremo, hemos sido testigos de numerosas denuncias a buscadores de internet que confundían a personas de raza negra con monos. Lo que

debemos analizar es por qué existe esta discriminación para poder prevenirla. Y existe básicamente porque el algoritmo se ha entrenado con imágenes de personas mayoritariamente de raza blanca. La profesora Lucía Benítez, en relación al estudio de los sesgos algorítmicos de género, apunta que “el origen del problema se encuentra en los *datasets* más utilizados para los *benchmarks* que miden el buen funcionamiento de los algoritmos de reconocimiento facial que son mayoritariamente masculinos (77%) y de origen caucásico (entre el 79 y el 86% de rostros blancos)” (Benítez 2019, 1313).

Debemos analizar factores como la publicidad, transparencia, trazabilidad, explicabilidad y auditabilidad de estos algoritmos para evitar incurrir en la vulneración de derechos fundamentales como la igualdad, privacidad, el derecho a la no discriminación o principios como la dignidad humana y el libre desarrollo de la persona (Simón 2021, 215-216).

4.2. Falta de transparencia y explicabilidad de los algoritmos

Otro factor que puede influir en la discriminación algorítmica es la falta de transparencia y explicabilidad de los algoritmos. En algunos casos, los algoritmos pueden ser “cajas negras” en el sentido de que no se puede entender cómo se han tomado ciertas decisiones. Si no se comprende cómo se han tomado estas decisiones, es difícil detectar y corregir cualquier discriminación que pueda haber ocurrido.

Los artículos 13 y 14 del RGPD (apartados 2.f y 2.d) desarrollan dos situaciones en las que es obligatorio informar al interesado sobre las decisiones automatizadas. Esto implica que cuando los datos se obtengan de los propios interesados o de terceros, siempre y cuando no hayan sido informados previamente o cuando exista una excepción a la obligación de informar, deberá indicarse la existencia de las decisiones automatizadas (incluida la elaboración de perfiles) y, en particular, la información sobre la lógica aplicada y las consecuencias previstas de dicho tratamiento para el interesado.

4.3. Falta de diversidad en el desarrollo de algoritmos

También la falta de diversidad en el desarrollo de algoritmos puede contribuir a la discriminación algorítmica. Si los equipos de desarrollo de algoritmos están compuestos principalmente por personas de un

solo grupo demográfico, es menos probable que se tengan en cuenta las perspectivas y experiencias de otros grupos. Esto puede llevar a que se desarrollen algoritmos que no funcionen bien para ciertos grupos o que sean discriminatorios hacia ellos.

A este respecto, los investigadores Pablo J. Boczkowski y Eugenia Mitchelstein advierten que "solo el 10% de la población no tiene acceso a internet en Estados Unidos. Ese porcentaje desciende al 89% entre los blancos, pero aumenta al 14% entre los latinos y al 15% entre las personas negras. Además, el 72% de los blancos tiene banda ancha en casa, en comparación con el 47% y el 57% entre latinos y negros, respectivamente. Esta disparidad está relacionada con la mano de obra del sector tecnológico digital de la economía: los latinos representan solo el 7% de los programadores, y los negros solo el 8%. Esta composición ha cambiado poco en los últimos quince años, a pesar del notable crecimiento de este sector durante el mismo periodo. El sueño de una tecnología neutra, igualitaria y posracial que marcó los primeros años del entorno digital terminó siendo solo eso: un sueño" (Boczkowski y Mitchelstein 2022).

4.4. *Falta de transparencia y responsabilidad en el diseño y uso de los algoritmos*

La falta de transparencia y responsabilidad en el diseño y uso de los algoritmos es otro factor que puede contribuir a la discriminación algorítmica. A menudo, los algoritmos son diseñados y utilizados por empresas privadas o agencias gubernamentales sin una supervisión adecuada o una regulación clara. Esto puede resultar en la falta de transparencia en cuanto a cómo funcionan los algoritmos y cómo se toman las decisiones basadas en ellos. Además, puede ser difícil determinar la responsabilidad cuando se produce una discriminación algorítmica, ya que puede haber múltiples partes involucradas en el diseño y uso de los algoritmos.

El RGPD, en su artículo 25 formula el principio fundamental de la privacidad desde el diseño y por defecto, recogiendo así el cambio de tendencia de la Unión Europea para mutar las políticas reactivas seguidas hasta el momento en políticas proactivas, cuya filosofía se centre más en la prevención y no en la sanción. En este sentido, el profesor Castellano (2021, 207) se refiere a la jurisdiccionalidad desde el diseño y por defecto.

Como acertada y concisamente señala el profesor Cotino (2019, 35), "para una IA *made in Europe* un principio clave será el de 'ética

por diseño' según el cual los principios éticos y legales, conforme al Reglamento General de Protección de Datos, el cumplimiento de la ley de competencia y la ausencia de sesgos de datos se implementan desde el inicio del proceso de diseño" (Comisión Europea 2018b: 5-9 y 19-21). Y a ello debe añadirse la exigencia de la transparencia y explicabilidad también en el diseño mismo. También, la Conferencia Internacional de Protección de Datos afirma un "enfoque general de 'ética por diseño' (ICDPPC 2018: ap. 4º)" (Cotino 2019, 36).

Otro de los factores fundamentales para fomentar la transparencia, y con ello la explicabilidad y confiabilidad de los algoritmos, es el hecho de poder acceder a su código fuente. Y este es hoy en día el caballo de batalla de asociaciones como CIVIO y de diferentes juristas que intentan que la Administración haga visible el mismo, aunque, de momento una Sentencia del Juzgado Central de lo Contencioso Administrativo¹ ha desestimado esta petición en relación al programa Bosco que determina quién es merecedor del bono social en los servicios eléctricos. Se ha demostrado que este algoritmo ha dejado fuera del sistema a personas que sí reunían los requisitos para pedir esta ayuda. Estamos plenamente de acuerdo con Javier de la Cueva cuando afirma: "La cuestión de hecho que constituye la base argumentativa de la demanda de Civio, que los técnicos informáticos están de facto escribiendo regulación jurídica, es algo que ya es indiscutible, sin que se trate de la mera teoría de un jurista estudioso del funcionamiento de las Tecnologías de la Información y Comunicación (TIC) en su nivel menos abstracto, más atómico y nuclear, que es el del código fuente" (Cueva 2019). Y, siguiendo a la profesora Gutiérrez David, podemos concluir que "cualquier regulación que se realice del uso de la IA en el ámbito público debería garantizar la transparencia de la actividad algorítmica, en sus dos vertientes, de publicidad proactiva y del derecho de acceso, con el fin de que los ciudadanos puedan saber y comprender: (i) en qué ámbitos concretos de la actividad pública se adoptan decisiones basadas total o parcialmente en tratamientos algorítmicos; (ii) en qué medida esas decisiones algorítmicas afectan a los derechos y libertades de los ciudadanos, individual o colectivamente; (iii) cómo y bajo qué criterios se han adoptado esa clase de decisiones por parte de las instituciones públicas; (iv) y, por su incidencia en la racionalidad y eficiencia del gasto público, cómo se han adquirido e implementado esos sistemas

¹ SAN 5863/2021 - ECLI:ES:AN:2021:5863: Audiencia Nacional. Juzgados Centrales de lo Contencioso. Madrid. Sección: 8. Fecha: 30/12/2021

de decisión algorítmica (ej. mediante desarrollos *in house* o mediante licitación), qué fines institucionales específicos se pretenden cumplir, qué necesidades públicas, satisfacer y por qué dicha implementación tecnológica innovadora es la mejor alternativa frente a otras soluciones” (Gutiérrez David 2021, 68).

5. Regulación de la discriminación algorítmica

5.1. Marco legal actual

La regulación de la discriminación algorítmica es un tema complejo y en constante evolución. Actualmente, existen diversas iniciativas legislativas y normativas a nivel internacional, regional y nacional que buscan abordar este problema.

A nivel internacional, la Organización para la Cooperación y el Desarrollo Económico (OCDE) ha emitido directrices sobre la inteligencia artificial que incluyen principios para evitar la discriminación y proteger los derechos humanos. Además, la Unión Europea (UE) ha publicado una estrategia sobre inteligencia artificial que incluye medidas para garantizar la ética y la transparencia en el uso de algoritmos.

En el ámbito regional, la Carta de Derechos Fundamentales de la UE prohíbe la discriminación por motivos de origen racial o étnico, y el Reglamento General de Protección de Datos (RGPD) establece el derecho a la explicación y la revisión de decisiones basadas en algoritmos.

A nivel nacional, algunos países han adoptado leyes específicas para abordar la discriminación algorítmica. Por ejemplo, en los Estados Unidos, la Ley de Derechos Civiles de 1964 ya prohibía la discriminación por motivos de raza, color, religión, sexo u origen nacional en el empleo, la educación y otras áreas. En Francia, la Ley de Orientación y Programación para la Reforzamiento de la Seguridad Interior (LOPPSI) incluye disposiciones para garantizar la transparencia y la evaluación de los sistemas de inteligencia artificial utilizados por diferentes organismos públicos. Asimismo, “la “Ley de la República Digital”, que entró en vigor en Francia en 2016, a partir de un debate sobre la transparencia. También, la *Association for Computing Machinery US Public Policy Council* (USACM, 2017) ha propuesto siete principios para el control de los algoritmos: la concienciación sobre los posibles sesgos presentes en sus procesos; la impugnación y compensación por el uso de los algoritmos; la responsabilidad de las

decisiones adoptadas a través de su uso; la transparencia sobre los procesos y decisiones, especialmente de cara al sector público; el control de la captación y procedencia de los datos con su revisión pública (aunque también figura una forma restringida a expertos); la posibilidad de auditar los resultados, y, por último, la validación y la prueba para conocer el impacto discriminatorio de sus fórmulas” y la Unión Estadounidense por las Libertades Civiles (*American Civil Liberties Union, ACLU*), ha lanzado una iniciativa relacionada con los sesgos algorítmicos: el proyecto *AI Now* (Benítez 2019, 1316).

A pesar de estos esfuerzos, la regulación de la discriminación algorítmica sigue siendo un desafío, ya que las tecnologías y los modelos cambian rápidamente y las leyes quedan obsoletas en un momento de continua mudanza. Es necesario un enfoque multidisciplinario y colaborativo que involucre a expertos en tecnología, derecho, ética y derechos humanos para abordar este problema de manera efectiva y proteger los derechos de los individuos.

Seguindo al profesor Ausín, podemos afirmar que “la reflexión ética sobre la IA constituye un instrumento privilegiado para el imprescindible empoderamiento tecnológico de la ciudadanía. Asimismo, y como necesario complemento de la reflexión ética, los poderes públicos y, en general, las instituciones y otras organizaciones, han de promover la formación y sensibilización de la sociedad sobre el potencial de la IA para generar valor económico y social, haciendo hincapié en los deberes de cuidado y protección mencionados anteriormente y, a la vez, en las oportunidades que un tratamiento correcto y ético de la IA ofrece para la investigación biomédica, la salud, la gestión administrativa, los servicios sociales, la atención a colectivos desfavorecidos, el desarrollo económico, la innovación...; esto es, para el bien común” (Ausín 2021, 14).

5.2. Necesidad de una regulación adecuada

La falta de regulación clara y adecuada en relación con la discriminación algorítmica ha sido un factor importante en la proliferación de prácticas discriminatorias en diversas áreas, incluyendo la contratación, la vivienda, la educación y el acceso a servicios públicos. La mayor parte de la normativa existente no ha sido diseñada para abordar este problema específico, y a menudo no ofrece soluciones adecuadas para proteger a las víctimas de la discriminación algorítmica.

Por lo tanto, resulta necesario implementar una regulación adecuada para abordar la discriminación algorítmica y proteger los

derechos humanos, especialmente los de los grupos más vulnerables. Se necesitan medidas legales claras que establezcan responsabilidades y obligaciones para los diseñadores y usuarios de algoritmos, y que garanticen la transparencia y la rendición de cuentas en su diseño y uso. Además, es importante considerar la necesidad de supervisión y sanciones para asegurar el cumplimiento de las regulaciones y la protección de los derechos fundamentales.

En el ámbito regulativo de la Unión Europea se está tratando de ofrecer respuestas adecuadas para prevenir futuras vulneraciones de los derechos fundamentales. Así, con respecto al tratamiento automatizado de datos, el artículo 22 del RGPD establece que “todo interesado tendrá derecho a no ser objeto de una decisión basada únicamente en el tratamiento automatizado, incluida la elaboración de perfiles, que produzca efectos jurídicos en él o le afecte significativamente de modo similar”. Además, se considera más ampliamente la cuestión de la utilización de algoritmos e inteligencia artificial en la Resolución del Parlamento Europeo del 16 de febrero de 2017 con recomendaciones para la Comisión sobre “normas de Derecho civil sobre robótica”; en la Comunicación de la Comisión Europea del 25 de abril de 2018 sobre “Inteligencia Artificial para Europa” (COM(2018) 237 final); en la Comunicación de la Comisión Bruselas “Generar confianza en la inteligencia artificial centrada en el ser humano” del 8 de abril de 2019, COM(2019) 168 final y, en el Libro Blanco sobre Inteligencia Artificial del 19 de febrero de 2020 (COM(2020) 65 final). En el marco del Consejo de Europa, el Comité de Expertos en Intermediarios de Internet publicó el estudio:

Algoritmos y Derechos Humanos en marzo de 2018 y poco después, el 3 de diciembre de 2018, la Comisión Europea para la Eficiencia de los Sistemas de Justicia (CEPEJ) aprobó la “Carta ética sobre la utilización de la inteligencia artificial en los sistemas de justicia y su entorno”.

En nuestro país, la transparencia algorítmica se exige constitucionalmente (arts. 9.3, y 24 CE) y por motivaciones legales generales, así como por algunas regulaciones específicas aplicables (art. 41 Ley 40/2015, arts. 13 y 11 artículo 13 Real Decreto 203/2021). Más positivas son regulaciones novedosas como la avanzada Ley 1/2022 valenciana de transparencia o el artículo 23 Ley 15/2022 inspirada en el apartado XVIII de la Carta de Derechos Digitales, a la que también se presta singular atención. En todo caso, aunque no está aprobado, se tiene especialmente en cuenta la transparencia exigida en razón del futuro Reglamento de IA de la UE (*Artificial Intelligence Act*) (Cotino 2023a, 22).

5.3. Principios de una regulación justa y equitativa

En orden a lograr una regulación de la discriminación algorítmica justa y equitativa, es necesario establecer algunos principios fundamentales. En primer lugar, debe garantizar el derecho a la no discriminación y la igualdad de trato para todas las personas, sin importar su origen, raza, género u otras características protegidas. “Los estereotipos delictivos asociados a diferentes grupos raciales pueden aumentar la idea de culpabilidad, especialmente cuando el origen étnico de los sospechosos es congruente con los estereotipos de la delincuencia, lo que da lugar a una investigación impulsada por esta errónea consideración y por la búsqueda de más pruebas confirmatorias (...) Estos estereotipos y muchos otros no son independientes, si no que interactúan entre sí, lo que da lugar a diferentes tipos de tratamientos y sentencias en función del prejuicio judicial: por ejemplo, no es sorprendente que los varones negros sean condenados a penas de reclusión más largas que cualquier otra combinación de raza y género” (Cippitani 2021, 5-6).

En segundo lugar, debe promover la transparencia y responsabilidad en el diseño y uso de los algoritmos, de manera que se puedan identificar y corregir posibles sesgos y discriminación. En tercer lugar, debe fomentar la participación y colaboración de todas las partes interesadas, incluyendo a la sociedad civil, los expertos en derechos humanos, los usuarios de los algoritmos y los desarrolladores. Finalmente, debe garantizar un enfoque basado en los derechos humanos, en el que se priorice la protección y promoción de los derechos fundamentales en todo momento.

Diversos documentos de la UE hacen hincapié sobre la importancia de preservar los derechos humanos frente a las injerencias de la inteligencia artificial. Así lo constatamos en la Resolución del Parlamento sobre la robótica, en su párrafo décimo, que se refiere especialmente a “la seguridad y la salud humana; la libertad, la intimidad, la integridad y la dignidad; la autodeterminación y la no discriminación, y la protección de los datos personales”. Se trataría de regular los algoritmos garantizando la confianza depositada por los usuarios. Estamos de acuerdo con el profesor Cotino cuando afirma que “Sin perjuicio de la necesidad de sólidos presupuestos éticos, jurídicamente punto de partida deben ser los principios comunes del Derecho constitucional y en particular los derechos y libertades de las sociedades democráticas. Ellos no sólo quedan afectados por el *big data* y la inteligencia artificial, sino que son principios básicos que han

de orientar las respuestas futuras. Ahora bien, puede ser necesaria una readecuación de tales principios” (Cotino 2017, 137).

Con motivo de la muy reciente sentencia del Tribunal Constitucional alemán², el profesor Cotino nos ofrece un exhaustivo análisis de esta, mostrándonos las claves de la futura regulación que debería hacerse extensiva a todo ciudadano europeo. Así, afirma que “la sentencia del Tribunal Constitucional alemán de 2023 ha anulado dos leyes que permitían la evaluación automatizada de datos con fines policiales. Además, establece elevados estándares de calidad de la ley para los tratamientos automatizados de datos para la prevención del delito. Asimismo, excluye el uso de sistemas IA. La regulación legal debe establecer los mínimos técnicos y organizativos y limitar los métodos y los datos utilizados. Esta importante sentencia se suma a una tendencia jurisprudencial exigente en Europa y España. Dadas las graves carencias legales en España, es imperativa la actuación de un legislador de calidad, incluso con el futuro Reglamento de la Unión Europea” (Cotino 2023b, 1).

6. Soluciones técnicas a la discriminación algorítmica

6.1. Técnicas de inteligencia artificial para la eliminación de sesgos

Las técnicas de inteligencia artificial pueden ser utilizadas para la eliminación de sesgos en los algoritmos. Una técnica comúnmente utilizada es la “reducción de sesgo”, que implica la modificación de los datos de entrenamiento o los modelos para reducir los mismos. Otras técnicas incluyen la selección de características y la regularización. Estas técnicas pueden mejorar la precisión y la equidad de los algoritmos, lo que a su vez puede reducir la discriminación algorítmica.

A modo de ejemplo, un equipo de investigadores del Instituto de Tecnología Massachusetts (MIT) ha desarrollado un algoritmo que apunta a reducir el sesgo racial y de género en los sistemas de reconocimiento facial basados en inteligencia artificial (IA). Según los resultados de sus pruebas, este algoritmo ha logrado reducir hasta en un 60% el sesgo que resulta del entrenamiento del reconocimiento facial de la IA en términos de género y raza. También se han

² Sentencia de 16 de febrero de 2023 (1 BvR 1547/19, 1 BvR 2634/20) de la primera cámara del Tribunal Constitucional Federal: https://www.bundesverfassungsgericht.de/er/20230216_1bvr154719.html

desarrollado técnicas para reducir el sesgo en los algoritmos de contratación y en los sistemas de evaluación crediticia.

En otro sentido, podemos destacar las políticas llevadas a cabo en Colombia, usando la minería de datos para su implementación en la sociedad, educación o salud tratando de sostener un pacto por la equidad. Como señalan los investigadores Enrique Arrieta y Amanda Vargas, “la analítica de datos aplicada a los documentos de políticas públicas, entre otros documentos gubernamentales, ofrece una poderosa herramienta a la sociedad, la cual puede ser utilizada para comprender dichas políticas sin sesgos ideológicos. De igual manera, la herramienta es de suma utilidad para los gobernantes, ya que a través de ella pueden verificar si sus discursos públicos son coherentes con los documentos legales que los describen, ya que son los documentos los que realmente determinan la hoja de ruta para el desarrollo del país” (Arrieta y Vargas 2021, 219).

Sin embargo, es importante tener en cuenta que estas técnicas no son una solución completa al problema de la discriminación algorítmica y deben ser utilizadas en conjunto con otras medidas, como la transparencia y la responsabilidad en el diseño y uso de los algoritmos.

6.2. *Evaluación de la calidad y equidad de los algoritmos*

La evaluación de la calidad y equidad de los algoritmos es un aspecto clave en la lucha contra la discriminación algorítmica. La falta de transparencia en el diseño y uso de los algoritmos puede llevar a la ocurrencia de sesgos y discriminación, lo que hace necesario implementar medidas de evaluación que permitan identificar y corregir estas prácticas.

Existen diversas técnicas y herramientas que permiten evaluar la calidad y equidad de los algoritmos. Entre ellas se incluyen los estudios de impacto y las auditorías algorítmicas. Los estudios de impacto son evaluaciones exhaustivas que buscan medir el impacto de los algoritmos en los derechos humanos y la equidad, mientras que las auditorías algorítmicas se centran en la evaluación de la calidad de los algoritmos y la identificación de sesgos y discriminación.

Es un hecho incontestable que los algoritmos que manejan datos que pueden afectar a derechos civiles, penales, sanitarios, crediticios o sociales deben ser auditables. Y tal y como pone de manifiesto el investigador del CSIC Txetxu Ausín (2021, 9), “para conciliar este principio con la protección de la innovación científica e industrial y la propiedad intelectual, debiera considerarse la atribución de

competencias de fiscalización a autoridades independientes. Más aún, cuando se pongan en juego valores fundamentales como los mencionados (salud, vigilancia policial, sistema penal, ayuda social...), los sistemas deberían ser trazables, esto es, se podrá identificar el proceso de decisión, las personas implicadas y las consecuencias que se deriven”.

La implementación de estas medidas de evaluación es crucial para garantizar la equidad y la justicia en el uso de los algoritmos, y para prevenir la discriminación algorítmica en todos los ámbitos, desde la selección de personal hasta la justicia penal y la atención médica.

7. Caso de estudio: Discriminación de los inmigrantes

7.1. Impacto de la discriminación algorítmica en los inmigrantes

La discriminación algorítmica puede tener un impacto especialmente negativo en los inmigrantes. En algunos casos, los sistemas de IA utilizados para procesos de selección de empleo y evaluación de crédito pueden considerar el origen nacional de una persona como un factor de riesgo, lo que puede llevar a una discriminación injusta. Además, algunos sistemas de vigilancia y control de fronteras basados en la IA pueden utilizar perfiles de riesgo que se basan en prejuicios y estereotipos, lo que puede llevar a la identificación errónea y detención de personas inocentes. Por lo tanto, como ya hemos expuesto anteriormente, es fundamental abordar la discriminación algorítmica en relación con los inmigrantes y desarrollar soluciones efectivas y sostenibles para garantizar su igualdad y no discriminación.

7.2. Ejemplos de discriminación algorítmica en el contexto migratorio

Además de los casos estudiados en las páginas precedentes (II. Discriminación algorítmica y derechos humanos: c) Ejemplos de discriminación algorítmica), vamos a tratar de analizar algunos más que afectan directamente a los inmigrantes. Uno de los ejemplos más conocidos de esta discriminación algorítmica se dio en el “sistema de puntos” de Canadá, utilizado para evaluar la elegibilidad de los solicitantes de inmigración. En este sistema, se otorgaban puntos a los solicitantes en función de criterios como la educación, la edad y la experiencia laboral. Sin embargo, se encontró que el sistema tenía sesgos en su diseño que discriminaban a los solicitantes de ciertos países, lo que llevó a una demanda colectiva en 2019.

Otro ejemplo es el sistema de *Minority Report* utilizado en la ciudad de Nueva York para predecir crímenes y evaluar el riesgo de reincidencia en los delincuentes. Este sistema se basaba en datos históricos que reflejaban la discriminación racial y de género en la aplicación de la ley, lo que llevó a que el sistema continuara perpetuando esta discriminación en lugar de corregirla. Si el algoritmo refleja que el barrio es peligroso se enviarán más policías, se detendrá por tanto a más personas y esto hace que se alimente al algoritmo porque parece más probable que en ese lugar haya más delincuentes. Sin embargo, los algoritmos únicamente hacen predicciones en base a análisis estadísticos sesgados en origen.

No debemos bajar la guardia porque, como advierte el profesor Cotino, "en EEUU la clave jurídica para admitir que los algoritmos y el big data incluyan información racial reside en que el algoritmo no haya elegido el factor raza con finalidad discriminatoria ("*because of*" "*discriminatory purpose*"), que es lo que prohíbe la doctrina *McCleskey v. Kemp* 481 U.S. 279, 291-92 (1987). Por el contrario, quizá se podría de admitir que se tome en cuenta el factor raza por ser un elemento que aumenta la precisión "a pesar de" (*in spite of*) sus efectos sobre un grupo identificable. Habrá que analizar rigurosamente cada caso concreto el funcionamiento y resultados de la IA, eso sí, bajo la presunción de discriminación" (Cotino 2023c, 302-303).

De este modo, podemos clasificar estas herramientas informáticas en función de los derechos afectados de los inmigrantes:

- Algoritmos de análisis de riesgos crediticios que pueden discriminar a los inmigrantes que no tienen historial crediticio en el país de destino, lo que dificulta su acceso a préstamos y servicios financieros.
- Algoritmos de contratación que pueden discriminar a los inmigrantes por su origen étnico o nacionalidad, incluso si tienen las habilidades y calificaciones necesarias para el trabajo.
- Algoritmos utilizados en el procesamiento de solicitudes de asilo que pueden tomar decisiones erróneas basadas en la nacionalidad o raza del solicitante, lo que resulta en un mayor riesgo de deportación o negación de asilo.
- Algoritmos de vigilancia y control fronterizo que pueden identificar erróneamente a los inmigrantes como una amenaza o peligro potencial, lo que puede llevar a detenciones injustas y deportaciones forzosas.

Ejemplos que acreditan la importancia de evaluar y corregir los sesgos en los algoritmos utilizados en la toma de decisiones que

afectan a los inmigrantes y otros grupos vulnerables. Como ha dejado escrito el profesor Esteban Guinea, podemos emplear la inteligencia artificial como una técnica de control social que puede derivar en resultados tan extremos como el del genocidio de los uigures (Esteban 2022, 390) Son muchas las instituciones que llevan tiempo advirtiendo de los abusos y vulneraciones de derechos humanos que se producen en los “campos de reeducación” chinos en la región de Xinjiang. Son frecuente las denuncias relativas a las esterilizaciones forzadas, violaciones, torturas sistemáticas y trabajos forzados, además de utilizar tecnología biométrica y de vigilancia para su rastreo.

Un ejemplo de posible discriminación algorítmica en España es el sistema SARA (Sistema Automatizado de Reconocimiento de la Atención), utilizado por algunas mutuas de trabajo en España para evaluar la validez de las bajas médicas de los trabajadores. Se ha cuestionado su objetividad y transparencia, ya que se desconoce la fórmula exacta que utiliza para calcular la probabilidad de fraude y algunos expertos han señalado que podría estar basado en factores irrelevantes como la edad, el género o el lugar de residencia del trabajador.

Una muestra que ilustra la discriminación algorítmica en España y sus implicaciones en los derechos humanos estaría reflejada en los siguientes casos:

Selección de candidatos: En 2019, se descubrió que el sistema utilizado por una empresa española para seleccionar candidatos para un programa de formación profesional tenía en cuenta criterios como la edad y el género, lo que resultaba en la discriminación de mujeres y personas mayores.

Seguros de coche: Las aseguradoras en España utilizan sistemas de puntuación para determinar el riesgo de cada conductor y, por lo tanto, el precio del seguro. Sin embargo, se ha informado de que estos sistemas pueden estar sesgados y resultar en una discriminación contra ciertos grupos, como jóvenes o personas de ciertos barrios.

Sistema de evaluación de profesores: En algunas comunidades autónomas en España, se utiliza un sistema de evaluación de profesores basado en indicadores numéricos y algoritmos. Se ha argumentado que este sistema es sesgado y puede conducir a la discriminación de ciertos profesores.

Vigilancia policial: El uso de cámaras de vigilancia por parte de la policía en España ha sido criticado por su potencial para la discriminación, especialmente contra ciertos grupos étnicos o sociales. Así, empresas como Renfe han tenido que revisar el software de vigilancia que utilizaban por hacer un mayor seguimiento de personas de otra raza o etnia.

8. Reflexiones finales

La discriminación algorítmica es un problema grave que afecta a los derechos humanos y puede perpetuar y amplificar las desigualdades existentes en nuestra sociedad. Es importante abordar este problema de manera efectiva, a través de la regulación y la evaluación de la calidad y equidad de los algoritmos, y de la implementación de técnicas de inteligencia artificial para la eliminación de sesgos.

Además, es necesario considerar el impacto que la discriminación algorítmica tiene en grupos vulnerables como los inmigrantes y trabajar en soluciones prácticas y sostenibles para garantizar la igualdad y no discriminación. Es fundamental seguir investigando y discutiendo sobre este tema para construir un futuro justo y equitativo en el uso de la tecnología y la inteligencia artificial.

Como reiteradamente hemos señalado, la discriminación algorítmica es un problema grave y urgente que puede tener perniciosas consecuencias en los derechos humanos, especialmente para grupos vulnerables como los inmigrantes. La falta de transparencia y responsabilidad en el diseño y uso de los algoritmos, así como la presencia de sesgos en los datos y modelos utilizados, son factores que contribuyen a la discriminación algorítmica. Como señala el profesor Serrano Antón centrándose en la Administración tributaria, pero siendo sus palabras extrapolables a otros ámbitos públicos y privados, "se deben extremar las medidas para una correcta ejecución de las tecnologías, alineando los conocimientos civiles, jurídicos y tecnológicos, junto con los aspectos éticos y buscando un equilibrio entre la privacidad a la que los ciudadanos tienen derecho y el cómo dar tratamiento a esos datos sin menoscabar su seguridad, ni crear situaciones discriminatorias, como consecuencia del diseño de los algoritmos. Resulta fundamental que las Administraciones tributarias que utilizan algoritmos e IA, lo lleven a cabo con el respeto de los derechos y garantías de los contribuyentes, siendo importante destacar que los algoritmos siempre deberían estar sujetos a revisión judicial" (Serrano 2021).

Aunque existen esfuerzos para regular y controlar la discriminación algorítmica, todavía queda mucho por hacer. La necesidad de una regulación justa y equitativa que tenga en cuenta los principios de igualdad y no discriminación es esencial. También es importante desarrollar técnicas de inteligencia artificial para eliminar los sesgos y garantizar la equidad y la calidad de los algoritmos.

Es importante seguir investigando y monitoreando la discriminación algorítmica y sus efectos en diferentes grupos y contextos, incluyendo

el contexto español. Solo a través de un enfoque colaborativo y multidisciplinario podemos abordar eficazmente este problema y promover una sociedad más justa e igualitaria para todos.

A modo de conclusión, la discriminación algorítmica es un problema grave que requiere atención y acción por parte de los responsables políticos, las empresas tecnológicas y la sociedad en general. Los algoritmos se encuentran cada vez más omnipresentes y tienen consecuencias importantes para los derechos humanos, especialmente para aquellos grupos que ya enfrentan discriminación en la sociedad.

Es necesario abordar este problema de manera efectiva y desarrollar soluciones prácticas y sostenibles que promuevan la igualdad y la no discriminación en la era digital. Esto requiere una mayor transparencia y responsabilidad en el diseño y uso de los algoritmos, así como la promoción de la diversidad y la inclusión en la tecnología y en la industria de la tecnología.

La regulación adecuada también es fundamental para garantizar que los algoritmos sean justos y equitativos, y se necesitan principios claros y bien definidos para orientar la regulación. Además, es importante que se realice una evaluación rigurosa y continua de la calidad y equidad de los algoritmos para garantizar que no estén causando daño a grupos vulnerables.

Como señala el profesor Cotino, en relación con la doctrina que se desprende de la reciente sentencia del Tribunal Constitucional alemán, en un importante artículo publicado, cuyo título no deja lugar a dudas respecto a su postura (“Una regulación legal y de calidad para los análisis automatizados de datos o con inteligencia artificial. Los altos estándares que exigen el Tribunal Constitucional alemán y otros tribunales, que no se cumplen ni de lejos en España”). “Pese a que puedan considerarse muy rigurosos estos criterios, pueden tomarse como punto de partida para hacer propuestas regulatorias. Ahora bien, creo que no debe orillarse o directamente prohibirse la regulación del uso de IA y los sistemas autónomos como hace sin excesiva justificación el TC alemán” (Cotino 2023b, 21). La insuficiencia de regulación legal en España no sólo se da respecto de tratamientos automatizados en el ámbito penal y judicial, sino también en el ámbito tributario.

Por último, apuntar la respuesta de la propia herramienta ChatGPT ante la cuestión de si cree que tiene sesgos en su algoritmo que podrían discriminar a los inmigrantes, y que nos da un reflejo fiel de lo expresado hasta ahora en cuanto al origen de los datos que sirven de base a la inteligencia artificial y la necesidad de evitar —técnica y

jurídicamente— esos sesgos: “ChatGPT: Como modelo de lenguaje basado en GPT-3.5, estoy diseñado para tratar de evitar sesgos y prejuicios en mis respuestas. Sin embargo, debido a que me he entrenado en una amplia variedad de datos tomados de Internet, existe la posibilidad de que haya absorbido ciertos sesgos presentes en esos datos. OpenAI, el equipo responsable de desarrollar y mantener ChatGPT, ha hecho esfuerzos para minimizar los sesgos, pero no es perfecto y pueden existir algunos sesgos residuales en mis respuestas. Es importante tener en cuenta que soy un modelo de inteligencia artificial y no tengo opiniones ni creencias propias. Mis respuestas se basan en el análisis de grandes cantidades de texto y en el intento de generar respuestas coherentes y útiles. No estoy programado para discriminar o mostrar prejuicios hacia ningún grupo de personas, incluyendo a los inmigrantes. Sin embargo, si notas alguna respuesta que parezca sesgada o inapropiada, te agradecería que me lo indiques. Esto ayudará a mejorar y refinar el modelo en el futuro. La eliminación de sesgos en la inteligencia artificial es un objetivo en constante evolución y es importante abordarlo de manera responsable y colaborativa” (ChatGPT mayo de 2023).

Esperamos que esta investigación sirva como un llamado a la acción y una fuente de información y reflexión sobre la discriminación algorítmica y sus impactos en los derechos humanos. Debemos trabajar juntos para asegurar que la tecnología sirva al bien común y promueva la igualdad y la justicia para todos.

9. Recomendaciones para la regulación y uso justo de los algoritmos

A continuación, se presentan algunas recomendaciones para la regulación y uso justo de los algoritmos:

Primera. Transparencia y responsabilidad: es necesario que los algoritmos sean transparentes en cuanto a su diseño y funcionamiento, y que los desarrolladores sean responsables de las decisiones que toman los algoritmos. Se debe exigir a las empresas y organizaciones que utilicen algoritmos que proporcionen información clara y comprensible sobre cómo funcionan sus sistemas, qué datos utilizan y con qué fines. También se debe establecer la responsabilidad correspondiente en caso de que se produzcan daños por el uso de algoritmos discriminatorios.

Segunda. Eliminación de sesgos: se deben utilizar técnicas de inteligencia artificial para reducir o eliminar los sesgos en los datos y

modelos utilizados en los algoritmos. Es importante realizar una evaluación constante de los algoritmos para detectar y corregir posibles sesgos. Para ello, se pueden utilizar técnicas de auditoría algorítmica y pruebas de impacto en los derechos humanos. Se debe realizar una evaluación continua de la calidad y equidad de los algoritmos, con el fin de identificar y corregir posibles discriminaciones.

Tercera. Participación de múltiples actores: se debe involucrar a múltiples actores, incluyendo a las comunidades afectadas por los algoritmos, en el diseño y evaluación de los mismos.

Cuarta. Regulación adecuada: es necesario que existan leyes y regulaciones adecuadas que garanticen la no discriminación algorítmica y promuevan el uso justo de los algoritmos. Se debe establecer una regulación dirigida a garantizar que los algoritmos no sean discriminatorios y respeten los derechos humanos. Esta regulación debe ser flexible y adaptarse a la evolución constante de la tecnología.

Quinta. Formación y educación: se debe promover la formación y educación sobre la discriminación algorítmica y su impacto en los derechos humanos, tanto entre los desarrolladores de algoritmos como entre los usuarios finales.

Referencias bibliográficas

- Aguado i Cudolà, Vicenç, Vera Parisio y Oscar Casanovas i Ibàñez. 2018. *El derecho a la ciudad: el reto de las smart cities*. Barcelona: Atelier. Acceso el 15 de mayo de 2022: <https://ucm.on.worldcat.org/oclc/1346587615>.
- Arrieta, Enrique y Amanda Vargas. 2021. «Analítica de datos aplicada al Plan Nacional de Desarrollo de 2018-2022, "Pacto por Colombia, Pacto por la equidad"». *Estudios Políticos* 62: 215-239. Doi.org/10.17533/udea.espo.n62a09.
- Asís, Rafael. 2020. «Inteligencia artificial, diversidad y derechos», *Uc3m Working paper*, Materiales de Filosofía del derecho, 20-04. Universidad Carlos III de Madrid.
- Ausín, Txetxu. 2021. «¿Por qué ética para la Inteligencia Artificial? Lo viejo, lo nuevo y lo espurio». *Sociología y Tecnociencia* 11, n. 2. Acceso el 27 de mayo de 2022: <https://dialnet.unirioja.es/servlet/oaiart?codigo=8155412>.
- Barona, Silvia. 2021a. *Algoritmización del derecho y de la justicia: de la inteligencia artificial a la smart justice*. Valencia: Tirant lo Blanch. Acceso el 13 de marzo de 2021: <https://biblioteca-tirant-com.bucm.idm.oclc.org/cloudLibrary/ebook/info/9788413786667>.
- Barona, Silvia. 2021b. «Justicia algorítmica y neuroderecho: una mirada multidisciplinar». *Alternativa* 124. Acceso el 5 de Julio de 2021: <https://biblioteca-tirant-com.bucm.idm.oclc.org/cloudLibrary/ebook/info/9788413972022>.

- Benítez, Lucía. 2019. «Ethics and transparency for detection of gender bias in algorithms». *Estudios sobre el mensaje periodístico* 25, n. 3: 1307-1320. Acceso el 6 de enero de 2022: <https://revistas.ucm.es/index.php/ESMP/article/view/66989>.
- Blanco, Carlos G. 2022. «El tratamiento de los algoritmos de IA para evitar sesgos discriminatorios y la regulación jurídica de los robots y la IA». En *Los Derechos Humanos en la inteligencia artificial: su integración en los ODS de la Agenda 2030*, dirigido por Elena García-Antón, 235-252, Cizur Menor: Thomson Reuters Aranzadi.
- Boczkowski, Pablo J. y Eugenia Mitchelstein. 2022. *El entorno digital*. Madrid: Siglo XXI Editores.
- Campione, Roger. 2021. «Recopilar y vigilar: Algunas consideraciones filosófico- jurídicas sobre inteligencia artificial», *Sociología y Tecnociencia* 11, n. 2: 123-139. Acceso el 8 de febrero de 2013: <https://dialnet.unirioja.es/servlet/oaiart?codigo=8155418>.
- Cippitani, Roberto, Antonella Mirabile y Martina Onofri. 2021. «“Objetividad científica” y sesgos en la toma de decisiones jurídicas: los casos de genética forense y de algoritmos», *Revista Justicia y Derecho* 4, n. 2: 1-22. doi.org/10.32457/rjyd.v4i2.1440.
- Corrales, José M. 2022. «Derecho al trabajo y discriminación: las políticas públicas de empleo como instrumento de igualdad», en *Los derechos humanos en la inteligencia artificial: su integración en los ODS de la Agenda 2023*, dirigido por Elena García-Antón, 399-422, Cizur Menor: Thomson Reuters Aranzadi.
- Cotino, Lorenzo. 2017. «Big data e inteligencia artificial. una aproximación a su tratamiento jurídico desde los derechos fundamentales», *Dilemata* 24: 131-150. Acceso el 6 de marzo de 2019: <https://dialnet.unirioja.es/servlet/oaiart?codigo=6066829>.
- Cotino, Lorenzo. 2019. «Ética en el diseño para el desarrollo de una inteligencia artificial, robótica y big data confiables y su utilidad desde el Derecho», *Revista Catalana de Dret Públic* 58: 29-48. doi.org/10.2436/rcdp.i58.2019.3303.
- Cotino, Lorenzo. 2023a. «Una regulación legal y de calidad para los análisis automatizados de datos o con inteligencia artificial. Los altos estándares que exigen el Tribunal Constitucional alemán y otros tribunales, que no se cumplen ni de lejos en España», *Revista General de Derecho Administrativo* 63: 1-22. Acceso el 22 de agosto de 2023: <https://www.uv.es/cotino/publicaciones/publicadoRGDACotin.pdf>
- Cotino, Lorenzo. 2023b. «Qué concreta transparencia e información de algoritmos e inteligencia artificial es la debida», *Revista Española de la Transparencia* 16: 17-63. doi.org/10.51915/ret.272
- Cotino, Lorenzo. 2023c. «Discriminación, sesgos e igualdad de la inteligencia artificial en el sector público». En *Inteligencia artificial y sector público: retos, límites y medios*, dirigido por Eduardo Gamero, 257-351. Valencia: Tirant lo Blanch.

- Cueva, Javier. 2019. «El derecho a no ser gobernados mediante algoritmos secretos», *El Notario del siglo XXI* 87. Acceso el 17 de septiembre de 2020: <https://www.elnotario.es/index.php/hemeroteca/revista-87/9636-el-derecho-a-no-ser-gobernados-mediante-algoritmos-secretos>
- Danesi, Cecilia C. 2022. «Inteligencia artificial y metaverso: el ordenamiento jurídico frente a los sesgos algorítmicos y los estereotipos de género en la realidad virtual». En *Los derechos humanos en la inteligencia artificial. Su integración en los ODS de la Agenda 2030*, dirigido por Elena García-Antón, 253-274, Cizur Menor: Thomson Reuters Aranzadi.
- Eguíluz, Josu A. 2020. «Desafíos y retos que plantean las decisiones automatizadas y los perfilados para los derechos fundamentales», *Estudios de Deusto* 68, n. 2: 325-367: doi.org/10.18543/ed-68(2)-2020pp3 25-367
- Esteban, Francisco J. 2022. «La inteligencia artificial, inteligencia emocional y derechos humanos: las enseñanzas del caso del genocidio uigur». *Los Derechos Humanos en la inteligencia artificial: su integración en los ODS de la Agenda 2030*, dirigido por Elena García-Antón, 381-398, Cizur Menor: Thomson Reuters Aranzadi.
- Febles, Gustavo. 2012. «Falsos negativos en mamografía», *Revista de Imagenología*, Época II - Vol. XVI, n. 2 / Junio 2013: 17-22. Acceso el 9 de abril de 2018: https://www.webcir.org/revistavirtual/articulos/diciembre13/uruguay/uru_espanol_a.pdf
- Ferrante, Enzo. 2021. «Inteligencia Artificial y sesgos algorítmicos ¿por qué deberían importarnos?», *Nueva Sociedad*, 294: 27-36. Acceso el 25 de mayo de 2022: <https://biblat.unam.mx/hevila/Nuevasociedad/2021/no294/3.pdf>
- García León, Alejandra G., Lima Sánchez, Dania Nimbe y Lucía Abascal. 2023. «Revisión sobre aspectos bioéticos en las decisiones de triaje en la unidad de cuidados intensivos durante la pandemia por Covid-19», *Revista de Bioética y Derecho* 57: 193-206: doi.org/10.1344/rbd2023.57.37282
- García-Antón, Elena y Francisco García Pascual. 2022. *Los derechos humanos en la inteligencia artificial: su integración en los ODS de la Agenda 2030*. Cizur Menor: Aranzadi.
- Gutiérrez David, María Estrella. 2021. «Administraciones inteligentes y acceso al código fuente y los algoritmos públicos. Conjurando riesgos de cajas negras decisionales», *Derecom* 30: 143-228. Acceso el 6 de septiembre de 2022: <https://www.derecom.com/derecom/>
- Jiménez Schlegl, Pablo. 2021. «El aprendizaje en sistemas autónomos e inteligentes: visión general y sesgos de fuentes de datos», *Arbor* 197, n. 802: doi.org/10.3989/arbor.2021.802005
- López Flores, Luis M. 2022. «Colonialidad algorítmica: racialización y sexualización mecanizada en el capitalismo digital». *Teknokultura* 19, n. 2: 231-239. doi.org/10.5209/tekn.78681
- López Pinilla, Ana M^a. 2021. «Inteligencia Artificial en el proceso penal: análisis a la luz del Fiscal Watson». *Pensamiento Jurídico* 54: 147-164. Acceso 12 de octubre de 2022: <https://revistas.unal.edu.co/index.php/peju/article/view/96091>

- Martín López, Jorge. 2022. «Inteligencia Artificial, sesgos y no discriminación en el ámbito de la inspección tributaria», *Crónica Tributaria* 182: 51-89. Doi.org/10.47092/CT.22.1.2.
- Morales, Gabriela. 2023. «Problemática antropológica detrás de la discriminación generada a partir de los algoritmos de la Inteligencia Artificial», *Medicina y ética* 34, n. 2: 429-455. Doi.org/10.36105/mye.2023v34n2.04.
- Roa, Marcela del Pilar, Jesús E. Sanabria y Katherin Dinás. 2022. «Uso del algoritmo compas en el proceso penal y los riesgos a los derechos humanos», *Revista Brasileira De Direito Processual Penal* 8, n. 1: 275-310. Doi.org/10.22197/rbdpp.v8i1.615
- Salas, Javier. 2019. «El algoritmo que discrimina a los pacientes negros sin conocer su raza», *El País*, 24 de octubre. Acceso el 24 de octubre de 2019: https://elpais.com/elpais/2019/10/24/ciencia/1571909798_596622.html
- Serrano, Fernando. 2021. *Inteligencia artificial y administración tributaria: eficiencia administrativa y defensa de los derechos de los contribuyentes*, Cizur Menor: Thomson Reuters Aranzadi.
- Simón, Pere. 2021. *Justicia cautelar e inteligencia artificial: la alternativa a los atávicos heurísticos judiciales*. Barcelona: Bosch.
- Sued, Gabriela E. 2022. «Culturas algorítmicas: conceptos y métodos para su estudio social», *Revista Mexicana de Ciencias Políticas y Sociales* 67, n. 246: 43-73. Doi.org/10.22201/fcpys.2448492xe.2022.246.78422.
- Valencia, Eduardo. 2018. *People analytics: analítica y datos para recursos humanos*. Cizur Menor: Aranzadi.
- Van Est, Rinie and Joost Gerritsen. 2017. *Human rights in the robot age: Challenges arising from the use of robotics, artificial intelligence, and virtual and augmented reality*, The Hague: Rathenau Instituut. Acceso el 10 de julio de 2020: <https://www.rathenau.nl/sites/default/files/2018-02/Human%20Rights%20in%20the%20Robot%20Age-Rathenau%20Instituut-2017.pdf>