INTELIGENCIA ARTIFICIAL Y POLITICAS PUBLICAS. EL JUICIO DE MAAT: LA PLUMA DE LOS DERECHOS FRENTE AL SESGO ALGORÍTMICO DISCRIMINATORIO, UNA CUESTIÓN DE EQUILIBRIO

INTELIGÊNCIA ARTIFICIAL E POLÍTICAS PÚBLICAS. O JULGAMENTO DO MAAT: A CANETA DOS DIREITOS CONTRA O VIÉM DISCRIMINATÓRIO ALGORÍTMICO, UMA QUESTÃO DE EQUILÍBRIO

María Méndez Rocasolano[1]

I. INTRODUCCIÓN

Como en la década de los años 90, cuando comencé a estudiar el medio ambiente desde el enfoque jurídico, la realidad mostraba los primeros atisbos de la necesaria participación del Derecho en el debate científico, igual ocurre con las nuevas tecnologías, especialmente con los algoritmos y la inteligencia artificial. Si entonces su contribución no sólo posibilitó el complejo entramado normativo e institucional dirigido a proteger y salvaguardar el capital natural de nuestro Planeta, también promueve y fomenta políticas ambientales en todos los países, lo que supuso el impulso de la conciencia ecológica, igualmente el Derecho hoy, tiene un papel capital en el contexto de la digitalización[2].

Ecología y digitalización se alían con la expresión "*ecotecnología"* para un futuro mejor y más justo. No en balde la Unión Europea los tiene consagrados como objetivos prioritarios en su política común, para lo que ha designado la mayor cantidad de dinero de la historia a

[1] Possui doutorado em Direito pela Universidad Complutense de Madrid(1999). Atualmente é Professor Assistente da Universidade Complutense de Madri, Professora da Universidad de Alcalá de Henares, Madrid, Consultora da Camara de Comercio de Madrid e Professor da Universidad Alfonso X El Sabio.

[2] El contenido de las siguientes consideraciones son fruto del Grupo de Investigación, Derecho Ecotecnología, Seguridad e Innovación, Claves de desarrollo del siglo XXI.Vid https://investigacion.ucam.edu/grupos/grupo/derecho-ecotecnologia,%20seguridad%20-e-innovacion/. El estudio de la materia ambiental desde la perspectiva elevada que proporciona el derecho constitucional, y el necesario estudio para la docencia de los aspectos legales y éticos de la Informática en la Facultad de Ingeniería de la computación, facilitan el entendimiento de conceptos basados en interacciones recogidas en procesos y su integración sistemática que aquí se vierten.

través del Mecanismo de Recuperación y Resiliencia en clave verde y digital con la intención de liderar estos retos[3].

Ciertamente queda un importante recorrido para lograr objetivos como la neutralidad climática o la implementación de la economía circular pero ha de reconocerse que en unas pocas décadas se ha producido un importante cambio de actitud, ahora falta la acción respecto al modelo productivo y de consumo. De hecho, así se denomina la llamada Agenda 20-30, "*la década de acción",* en la que se han comprometido casi todos los países del mundo a llevar a cabo soluciones sostenibles para enfrentarse a los principales desafíos del mundo, desde la pobreza y la igualdad de género, hasta el cambio climático, el cierre de la brecha financiera y la desigualdad [4].

Son múltiples las perspectivas desde las que podríamos hablar de nuevas tecnologías y Derecho, no todas concluyen con satisfacción, pues nos enfrentan a riesgos y preocupaciones inexistentes anteriormente. Probablemente se pensó lo mismo del fuego, la mecanización, la producción en masa, la electrónica o la energía nuclear, a las que debemos las tres primeras revoluciones industriales. El entorno digital, moldeado por algoritmos e IA, tiene el potencial de ser tan influyente en nuestras vidas como el mundo físico que nos rodea.

Desde el enfoque de la igualdad es crucial asegurar que este nuevo mundo digital esté construido sobre los cimientos de la equidad, la justicia y el respeto por la dignidad de la persona, pues el proceso de desarrollo de la civilización digital ha de ajustarse a la moral y la ética que reconoce la igualdad de todos los seres humanos sin razón de discriminación alguna por razón de sexo, religión o cualquier otra circunstancia social o económica.

Los aportes de otras civilizaciones incluso las lejanas en el tiempo, nutren nuestro presente, especialmente las del Derecho que como expresión cultural establece normas e instituciones para regular la conducta humana en sociedad, basado en postulados de Justicia. El

[3] Para una mayor información *Vid* https://commission.europa.eu/business-economy-euro/economic-recovery/recovery-and-resilience-facility_es

[4] Adelantándose a la década que comenzó en el 2020, el Secretario General de la ONU en 2019 indicó que para logra las pretensiones de los Objetivos del Desarrollo Sostenible eran necesarias acciónes en tres niveles: acciones a nivel mundial para garantizar un mayor liderazgo, más recursos y soluciones más inteligentes; acciones a nivel local que incluyan las transiciones necesarias en las políticas, los presupuestos, las instituciones y los marcos reguladores de los gobiernos, las ciudades y las autoridades locales; y acciones por parte de las personas, incluidas la juventud, la sociedad civil, los medios de comunicación, el sector privado, los sindicatos, los círculos académicos y otras partes interesadas, para generar un movimiento imparable que impulsara las transformaciones necesarias.

símbolo de la balanza con el que representamos universal e históricamente la Justicia ya estaba en la mitología egipcia donde el corazón del difunto era pesado contra la pluma de Maat, diosa de la verdad y la justicia, para determinar su destino en el más allá.

Esta representación del equilibrio entre el bien y el mal, la verdad y la falsedad, encuentra un eco sorprendente en nuestra era digital. A continuación utilizaré esta metáfora para reflexionar sobre como el Derecho proporciona el punto de equilibrio entre el potencial transformador de la Inteligencia Artificial y los derechos humanos fundamentales de igualdad. Con especial referencia a la igualdad de género.

El desafío que enfrentamos es monumental. Como advierte la experta en ética de la IA, Kate Crawford, "los sistemas de IA están creando una nueva era de optimización y control, a menudo a expensas de la autonomía humana y la autodeterminación"[5]. Este control algorítmico, cuando está sesgado, puede tener consecuencias devastadoras para los individuos y los grupos marginados, desequilibrando la balanza.

Igual que en el juicio de Maat, podemos restablecer dicho equilibrio a través de la "pluma de los derechos" que representa los valores fundamentales, los principios inalienables que definen nuestra humanidad y que se catalizan en la idea de la dignidad de la persona humana. Frente a ella, el "sesgo algorítmico" representa los peligros potenciales de una tecnología mal implementada o mal regulada.

En las páginas siguientes, exploraremos este moderno juicio de Maat. Examinaremos la naturaleza del sesgo algorítmico, su impacto en la dignidad humana y los derechos fundamentales por sus efectos discriminatorios. Reflexionaremos sobre, las consecuencias jurídicas y los medios que tiene el Derecho para garantizar la igualdad de todos los seres humanos como valor, como principio y como derecho fundamental y fundamentador de instituciones jurídicas[6].

[5] Crawford, K. (2021). Atlas of AI: Power, Politics, and the Planetary Costs of Artificial Intelligence. Yale University Press.

[6] El pensamiento jurídico, lleva tratando el concepto de igualdad desde la época clásica, son múltiples las aportaciones magistrales sobre la materia, recordemos aquí que primero fue el principio de igualdad ante la ley el que estableció que todos somos titulares de los mismos derechos, más tarde las revolución Francesa y la Revolución industrial estableció la igualdad de acuerdo con los principios de capacidad y talento, base del principio de no discriminación. Que se ha completado en la modernidad superando las diferencias sociales y económicas, así como las diferencias entre los sexos. En la actualidad su carácter prescriptivo resulta fundamental Ferrajoli, L. (2006). Derechos y Garantías. La ley del más débil. Madrid: Editorial Trotta, págs. 78-79. Esas tres dimensiones de la igualdad (como valor, como principio y como derecho) están bien destacadas por

Conversaremos con la IA y recordaremos algunos casos muy conocidos que ilustran estos desafíos, consideraremos el marco legal actual y sus limitaciones, y propondremos algunas ideas que pudieran utilizarse para el reconocimiento de nuevos derechos, al servicio de la denominada "Justicia Digital" para restaurar al menos el equilibrio entre el avance tecnológico y los derechos humanos.

Como señala Lawrence Lessig, "el código es la ley" en el ciberespacio[7] . En la era de la IA, debemos asegurarnos de que este código refleje nuestros valores más elevados y proteja los derechos de todos los individuos. Solo entonces podremos superar el juicio de Maat en este siglo XXI, donde ante la duda, la balanza debe siempre inclinarse a favor de la dignidad de la persona y los derechos que le son inherentes.

En la analogía del juicio de Maat, producto de la utilización de la inteligencia artificial en este asunto, he seleccionado algunos casos conocidos que resultan testimoniales y proporcionan evidencias concretas de cómo los sistemas de IA pueden afectar la vida de las personas y sus derechos. Conozcámoslos en primer lugar.

1.EL TESTIMONIO DE THOTH: CASOS REALES

Siguiendo con la mitología egipcia, Thoth era el dios de la sabiduría y el escriba que registraba el juicio de los muertos. Los escritos de su papiro daban fe de lo acontecido, igual que los siguientes casos que han puesto de manifiesto la realidad sobre la que se teorizaba al respecto del sesgo algorítmico discriminatorio. Su estudio ha enriquecido la doctrina respecto a los efectos discriminatorios del sesgo algorítmico relativo a las mujeres. Comencemos con el examen de casos reales, para luego analizar y buscar posibles soluciones.

1. COMPAS (Correctional Offender Management Profiling for Alternative Sanctions): Es un algoritmo utilizado en el sistema de justicia penal de Estados Unidos para evaluar el riesgo de reincidencia de los acusados.

Julia Angwin y sus colegas en ProPublica realizaron un análisis exhaustivo de este sistema y encontraron que: "El algoritmo falló de manera desproporcionada al predecir la reincidencia para acusados negros en comparación con acusados blancos,

Antonio PEREZ LUÑO (1987). «Sobre la igualdad en la Constitución española», en el Anuario de Filosofía del Derecho , vol IV, 133 y ss

[7] Lessig, L. (1999). Code and Other Laws of Cyberspace. Basic Books.

etiquetando erróneamente a los acusados negros como futuros criminales casi el doble de veces que a los acusados blancos"[8], lo que se hizo especialmente relevante para el caso de las mujeres.

2. Algoritmo de contratación de Amazon: La compañía desarrolló un sistema de IA para evaluar solicitudes de empleo, pero tuvo que abandonarlo cuando se descubrió que había una brecha de discriminación de género pues penalizaba los currículos que incluían la palabra 'mujeres'. Ifeoma Ajunwa argumenta que este caso demuestra cómo "los algoritmos de contratación pueden perpetuar y amplificar los sesgos históricos en un lugar de trabajo" [9], para lo que Miranda Bogen y Aaron Rieke sugieren "una supervisión cuidadosa y continua para evitar la discriminación algorítmica"[10].

3. Reconocimiento facial y sesgo racial:

En estado Unidos varios estudios han demostrado que los sistemas de reconocimiento facial tienen tasas de error más altas para personas de color, especialmente mujeres de piel oscura. Joy Buolamwini y Timnit Gebru encontraron que: "Los sistemas de clasificación de género tuvieron tasas de error de hasta el 34.7% para mujeres de piel oscura, en comparación con tasas de error de 0.8% para hombres de piel clara" [11].

4. Algoritmos para la concesión de créditos:

Lisa Rice y Deidre Swesnik argumentan que: "Los modelos de calificación crediticia, aunque aparentemente neutrales, a menudo utilizan factores que sirven como proxies para el género, la raza y el origen étnico, perpetuando así la discriminación histórica" [12].Mikella Hurley y Julius Adebayo proponen que "la regulación de los

[8] Angwin, J., Larson, J., Mattu, S., & Kirchner, L. (2016). Machine Bias. ProPublica.
[9] Ajunwa, I. (2020). The Paradox of Automation as Anti-Bias Intervention. Cardozo Law Review, 41(5), 1671-1742.
[10] Bogen, M., & Rieke, A. (2018). Help Wanted: An Examination of Hiring Algorithms, Equity, and Bias. Upturn.

[11] Buolamwini, J., & Gebru, T. (2018). Gender Shades: Intersectional Accuracy Disparities in Commercial Gender Classification. Proceedings of the 1st Conference on Fairness, Accountability and Transparency, 77-91.
[12] Rice, L., & Swesnik, D. (2013). Discriminatory Effects of Credit Scoring on Communities of Color. Suffolk University Law Review, 46, 935-966.

algoritmos de calificación crediticia debe ir más allá de la simple prohibición del uso explícito de características protegidas"[13] .

5. Algoritmos de precios dinámicos:

Arvind Narayanan demostró que: "Los precios mostrados a los usuarios de sitios de comercio electrónico pueden variar según su ubicación geográfica, historial de navegación y otros factores personales", como el ser mujer[14]. Maurice Stucke y Ariel Ezrachi sugieren que "la personalización de precios a través de algoritmos puede llevar a una forma de discriminación de precios de primer grado, planteando preocupaciones sobre la equidad y la privacidad del consumidor" [15].

El sesgo algorítmico no es simplemente un problema técnico, sino que tiene profundas implicaciones sociales y éticas. Como señala Frank Pasquale en "The Black Box Society", "los algoritmos no solo reflejan la realidad; le están dando forma activamente" [16]. En cualquier caso, que el sesgo algorítmico desconoce la igualdad de Trato y no discriminación, es una evidencia.

2.LA BALANZA DESEQUILIBRADA: LA DISCRIMINACIÓN DEL SESGO ALGORÍTMICO

Nadie duda hoy del especial protagonismo de los algoritmos que son utilizados para la toma de múltiples decisiones que afectan a nuestras vidas, desde la selección de productos y servicios que nos ofertan en las redes sociales, las múltiples aplicaciones que usamos en nuestra cotidianidad, como escuchar o ver películas desde Spotify o Netflix, hasta las cirugías que salvan nuestras vidas. Igualmente están desarrollándose para proveer mejores servicios públicos, para agilizar procedimientos o para facilitar la toma decisiones. A pesar de sus bondades se produce un desequilibrio que afecta derechos y obligaciones por sistemas algorítmicos con reducida transparencia, casi nulo conocimiento público, y escasos criterios de

[13] Hurley, M., & Adebayo, J. (2016). Credit Scoring in the Era of Big Data. Yale Journal of Law and Technology, 18(1), 148-216.
[14] Narayanan, A. (2013). Online Price Discrimination: Conspicuous by its Absence. 33 Bits of Entropy.
[15] Stucke, M. E., & Ezrachi, A. (2017). How Digital Assistants Can Harm Our Economy, Privacy, and Democracy. Berkeley Technology Law Journal, 32, 1239-1299.
[16] Pasquale, F. (2015). The Black Box Society: The Secret Algorithms That Control Money and Information. Harvard University Press.

responsabilidad. Todo ello repercute en el concepto de participación ciudadana consciente, elemento esencial en la determinación de lo que hasta hoy hemos entendido como Estado de Derecho y Democracia.

Es una cuestión no solo técnica, sino también filosófica y política, Russell, profesor de Ciencias de la Computación en Berkeley entiende que, "Debemos preguntarnos qué tipo de sociedad queremos construir con la IA y asegurarnos de que nuestros valores humanos guíen el desarrollo tecnológico."[17] La IA está transformando la forma de hacer política, la gestión pública y las dinámicas propias de la gobernanza pues como resultado concreto de la cuarta revolución tecnológica representa para la humanidad un cambio de paradigma cultural que transforma nuestra manera de relacionarnos con nosotros mismos, con los otros y con el medio, desde la aldea de Botswana hasta Nueva York.

Algunos lo consideran una magnífica oportunidad, otros creen que representa un riesgo existencial para la humanidad. Aunque la velocidad de gestión de información y por lo tanto la conclusión ante cualquier cuestión o tarea sea superior a la que las neuronas pueden procesar y nos parezca extraordinario que la inteligencia artificial resuelva muchas de las pretensiones, deseos y sueños de nuestros días, los algoritmos en cuanto creaciones humanas potencialmente contienen errores, y desde la perspectiva de la libertad y la igualdad, generan sesgos discriminatorios como acabamos de ver.

Estos sesgos algorítmicos ponen de manifiesto que los resultados de la IA ni son infalibles ni objetivos completamente. En palabras de O'Neil, autora de "*Weapons of Math Destruction"* el sesgo algorítmico no es solo un problema técnico, sino un reflejo de los retos sociales más amplios que enfrentamos. Son errores sistemáticos y repetibles en un sistema computacional que crean resultados injustos, como privilegiar a un grupo arbitrario de usuarios sobre otros. Y privilegian a los hombres frente a las mujeres. Es crucial entender que los algoritmos no son inherentemente objetivos, sino que pueden perpetuar y amplificar los sesgos existentes en nuestra sociedad[18]. En el mismo sentido Keyes advierte sobre los peligros de la

[17] Russell, S. (2019). Human Compatible: Artificial Intelligence and the Problem of Control. Viking.
[18] O'Neil, C. (2016). Weapons of Math Destruction: How Big Data Increases Inequality and Threatens Democracy. Crown Publishing Group.

"objetividad algorítmica", argumentando que "la creencia en la neutralidad de la tecnología puede ocultar y perpetuar sesgos perjudiciales"[19], lo que es todavía más peligroso.

Los orígenes del sesgo algorítmico se deben a que el algoritmo se entrena con datos históricos que a menudo reflejan prejuicios y discriminaciones pasadas según Safiya Umoja Noble, autora de "Algorithms of Oppression". Además, la poca participación de las mujeres en la informática supone falta de diversidad en los equipos de desarrollo, lo que en muchos casos llevan a "puntos ciegos y suposiciones no cuestionadas"[20] . Igualmente, la falta de criterios morales y éticos en la definición de estos objetivos nos lleva a resultados no deseados a través de las denominadas variables proxy. Estas variables aparentemente neutras, se ha observado que en realidad sirven como proxies para características protegidas como la raza o el género que afectan directamente las conquistas logradas sobre la igualdad de las personas, especialmente las de las personas de color y las mujeres.

La potencialidad disruptiva de la inteligencia artificial con el procesamiento del lenguaje natural representa un paradigma de aprendizaje automático, de forma que los algoritmos dan posibles respuestas a partir de un corpus formado por ejemplos y datos. En estos modelos de procesamiento del lenguaje natural, se tienden a asociar ciertos trabajos y características con géneros específicos, perpetuando los mencionados estereotipos de forma que todo el esfuerzo emancipatorio y de reconocimiento de la igualdad quiebra rotundamente. Así lo ponen de manifiesto Joy Buolamwini y Timnit Gebru, que en su estudio "Gender Shades", demuestran que los sistemas de reconocimiento facial comerciales tenían tasas de error significativamente más altas para mujeres de piel oscura[21].

En cuanto al sesgo socioeconómico discriminatorio, Virginia Eubanks, en su libro "Automating Inequality", documenta cómo los sistemas automatizados de toma de decisiones en servicios sociales a menudo perjudican desproporcionadamente a las comunidades de bajos ingresos y dentro de ellas especialmente de nuevo a las mujeres pues tienen ante el mismo trabajo menores salarios [22] . También la edad está sesgada, investigadores como Alex Rosenblat

[19] Keyes, O. (2018). The Misgendering Machines: Trans/HCI Implications of Automatic Gender Recognition. Proceedings of the ACM on Human-Computer Interaction, 2(CSCW), 1-22.
[20] Noble, S. U. (2018). Algorithms of Oppression: How Search Engines Reinforce Racism. NYU Press.
[21] Buolamwini, J., & Gebru, T. (2018). Gender Shades: Intersectional Accuracy Disparities in Commercial Gender Classification. Proceedings of the 1st Conference on Fairness, Accountability and Transparency, 77-91.
[22] Eubanks, V. (2018). Automating Inequality: How High-Tech Tools Profile, Police, and Punish the Poor. St. Martin's Press.

han encontrado que los algoritmos de publicidad en plataformas de redes sociales pueden discriminar por edad en anuncios de empleo[23].

Así las cosas, las mujeres de color y mayores son claramente discriminadas por errores del sesgo algorítmico en muchas situaciones en las que se utiliza la IA[24].

Someter el sesgo algorítmico a las reglas del reconocimiento de los derechos humanos representa uno de los desafíos más significativos en el desarrollo y despliegue de sistemas de Inteligencia Artificial al servicio de la humanidad, especialmente porque su procesamiento no siempre es transparente o auditable.

Tener en cuenta y entender la verdadera naturaleza de la IA, como herramienta únicamente y no como entidad decisional, es fundamental.

El organicismo que tanto ha ayudado a comprender conceptos complejos nos ayuda a la hora de comprender la IA. Pude decirse que replica el proceso de interacción neuronal humano y da soluciones razonadas a una problemática planteada. De Andrés[25] lo expresa brillantemente como una suerte de neuronas artificiales interconectadas entre sí que reciben información, la humana sinapsis neuronal se sustituye por una conexión de entrada y salida mediante un código binario (formado solamente por 0 o 1) o real con cualquier número desde el -∞ hasta el +∞, en el cual va encriptado un mensaje, que es enviado a todas aquellas unidades con las cuales tenga una conexión. La primera actividad sería la que se corresponde con procesos de aprendizaje o entrenamiento y la segunda es una función de transferencia.

Lo importante para el Derecho es qué hace, cómo se utiliza, qué derechos y obligaciones afecta o genera y cuáles son las consecuencias jurídicas de su acción u omisión.

[23] Rosenblat, A. (2018). Uberland: How Algorithms Are Rewriting the Rules of Work. University of California Press.

[24] No debemos olvidar el sesgo geográfico, diversos estudios han demostrado que los modelos de IA entrenados principalmente con datos de ciertas regiones geográficas pueden funcionar mal cuando se aplican a otros lugares, como lo demuestran los desafíos en la aplicación de modelos de diagnóstico médico desarrollados en países occidentales a poblaciones en África o Asia.

[25] De Andrés, T. (2002). Homo Cybersapiens. La inteligencia artificial y la humana. Pamplona, España, EUNSA, 90 , interesa también la lectura de ¿Qué tan natural es la inteligencia artificial? Sobre los límites y alcances de la biomímesis computacional de Héctor Velázquez Fernández, Colegio de Bioética de Nuevo León, México Naturaleza y Libertad. Número 12, 256.El conocimiento neurocientífico de la operatividad cerebral del ser humano opera desde la perspectiva de la tercera persona; mientras que lo más íntimo de nuestra operación mental, esa que solo es autotransparente al sujeto en el que se da, pertenecen a la perspectiva de la primera persona. En la dimensión de la primera persona, nos sabemos autoconscientes, autoreflexivos, autoorganizados en una biografía por la que advertimos que sabemos que existimos, sabemos quiénes somos y qué hacemos. Esto no está en el escenario de la IA porque para funcionar ésta no requiere de una singularidad biográfica del procesamiento acelerado de información que elabora. Y lo mismo puede decirse de los sistemas expertos y las redes neuronales artificiales.

Los analistas de la Inteligencia Artificial, coinciden en que representa la simulación de procesos de inteligencia humana como el aprendizaje, el razonamiento, y la autocorrección, de forma que le permite funcionar adecuadamente y con previsión en su entorno.

En el *Deep Learning*, el análisis de los datos suministrados determina la viabilidad y conveniencia de ejecutar una determinada operación, tomando así decisiones "razonadas" con base a la probabilidad de éxito y a una serie de variables. No es de estrañar que nos hayamos preguntado si es posible adjudicarles algún tipo de personalidad jurídica y por tanto derechos y responsabilidades. La Comisión de Asuntos Jurídicos del Parlamento Europeo, propuso en el año 2016[26] el concepto de "personalidad electrónica", aplicable "[…] a los supuestos en que los robots puedan tomar decisiones autónomas inteligentes o interactuar con terceros de forma independiente, sin embargo en el 2019 reconoció claramente la no necesidad de otorgar una personalidad jurídica a las tecnologías digitales emergentes, pues los derechos de la personalidad están vinculados esencialmente con la dignidad de la persona humana y la IA no lo es.

La IA está dirigida por el objetivo de su programación, carece de facultades cognitivas, de sentimientos y fundamentalmente de capacidad de adaptación para sobrevivir pues puede saber, pero no siente lo que es el hambre, el dolor o la muerte, tampoco la fé, el amor, la alegría, la esperanza o el deseo. Por todo ello no tiene sentido su categoría como persona[27] y por ende no pueden reconocérsele los derechos que les son inherentes, de entre ellos el primero y fundamental la vida que no puede atribuirse a un botón de encendido dependiente de una conexión eléctrica.

[26] La Comisión de Asuntos Jurídicos del Parlamento Europeo, señalaba en el 2016 que … "el propósito de reconocer dicha personalidad a los robots sería hacerles aptos para ser titulares de relaciones jurídicas con sus correspondientes derechos y obligaciones, y tener un cierto reconocimiento jurídico de su subjetividad, fundamentalmente en derechos de naturaleza patrimonial, pero no los constitucionales ni los de la personalidad, absolutamente consustanciales a la dignidad de los seres humanos.

[27] La cuestión está resuelta desde el Derecho de forma similar en todos los ordenamientos jurídicos del Mundo. Regulada normalmente en los códigos civiles con diferentes expresiones indican que para ser considerado persona desde el punto de vista legal hay simplemente que serlo. En este sentido la Comisión Mundial de Ética del Conocimiento Científico y la Tecnología de la Organización de las Naciones Unidas para la Educación, la Ciencia y la Cultura (2017), indica la necesidad de tener en cuenta el status moral por el que sería una contradicción el etiquetar a un sistema de inteligencia artificial bajo la denominación de "persona" pues carece de la capacidad de realizar deliberaciones racionales y morales que elige para dirigir su propia conducta a través de la voluntad. Las personas son fines en sí mismas y no pueden concebirse nunca de manera instrumental como simples medios para fines externos a ellas.

El debate queda abierto para los filósofos del Derecho. Sigamos ahora con la regulación que propone Europa, que orienta en cómo debemos de tratarla antes de sumergirnos en el examen entre el potencial transformador de la Inteligencia Artificial (IA) y la protección de los derechos humanos fundamentales, teniendo en cuenta que el sesgo algorítmico representa un desequilibrio fundamental en la balanza de la justicia digital en el caso del género. Como señala Kate Crawford, "la IA está creando nuevas asimetrías de poder"[28]. Abordar este desequilibrio requiere un enfoque multifacético que involucre a tecnólogos, legisladores, académicos y la sociedad civil. Solo a través de un esfuerzo concertado podremos esperar crear sistemas de IA que sean verdaderamente justos y equitativos

3.EL PESO DE LA DIGNIDAD HUMANA EN LA IGUALDAD DE GÉNERO

La mitad de la población humana son mujeres. Hablar de igualdad de género no es un tema aislado, es hablar de la dignidad humana, elemento fundamental y fundamentador de los derechos humanos y la ética. En el contexto del sesgo algorítmico y la Inteligencia Artificial, este concepto adquiere una nueva dimensión de complejidad y urgencia.

El filósofo Jürgen Habermas la define como "la fuente moral de la que todos los derechos fundamentales derivan su sustento"[29]. Martha Nussbaum, en su enfoque de las capacidades, argumenta que la dignidad humana requiere que cada individuo tenga la oportunidad de desarrollar ciertas capacidades centrales, incluyendo la razón práctica y la afiliación[30]. En el entorno digital, esto implica que los sistemas automatizados no deben menoscabar estas capacidades fundamentales que en el sesgo algorítmico impacta a través de la discriminación y trato desigual. Virginia Eubanks, en su libro "Automating Inequality" argumenta que la automatización de la desigualdad refuerza la idea de que algunas personas son menos valiosas que otras, lo que se observa en los sistemas algorítmicos de servicios

[28] Crawford, K. (2021). Atlas of AI: Power, Politics, and the Planetary Costs of Artificial Intelligence. Yale University Press.
[29] Habermas, J. (2010). The concept of human dignity and the realistic utopia of human rights. Metaphilosophy, 41(4), 464-480.
[30] Nussbaum, M. C. (2011). Creating capabilities: The human development approach. Harvard University Press.

sociales que pueden tratar a las mujeres especialmente en situación de pobreza como "sospechosas por defecto", socavando su dignidad [31].

También en la negación de oportunidades Cathy O'Neil describe cómo los algoritmos sesgados pueden reducir oportunidades de empleo, educación o créditicia a ciertos grupos, normalmente representados por mujeres limitando su capacidad para participar plenamente en la sociedad[32]. Esto no solo afecta las oportunidades de vida de las mujeres, sino que también erosiona su sentido de valor y pertenencia a la comunidad.

Como consecuencia del error que representan, los sesgos algorítmicos refuerzan estereotipos, reflejan y amplifican nuestros prejuicios sociales más profundos como demuestra Safiya Umoja Noble, en "Algorithms of Oppression" explicando que los motores de búsqueda afectan la autopercepción y la dignidad de grupos marginados [33], estableciéndose brechas indignas en la igualdad de género.

El impacto del sesgo algorítmico en la dignidad humana tiene la cara de una mujer, de color, gitana, sudamericana o perteneciente a alguna minoría étnica, mayor de 45 años y con alguna discapacidad. Es un impacto profundo, multifacético pues afecta no sólo a los derechos de igualdad sino al resto de los derechos humanos con los que está interconectados. En este sentido Frank Pasquale advierte sobre el peligro de la "sociedad de la caja negra", donde las decisiones algorítmicas opacas pueden privar a las personas de su autonomía y capacidad de autodeterminación,[34] lo que plantea preguntas fundamentales sobre la participación del criterio humano en una era de toma de decisiones automatizada. Shoshana Zuboff, en "The Age of Surveillance Capitalism", argumenta que la omnipresencia de los sistemas algorítmicos puede llevar a una sensación de impotencia y resignación entre los individuos[35]. Se advierte el peligro de un "futuro sin libertad de elección" donde la Democracia y el Estado de Derecho quedan afectos.

[31] Eubanks, V. (2018). Automating Inequality: How High-Tech Tools Profile, Police, and Punish the Poor. St. Martin's Press.
[32] O'Neil, C. (2016). Weapons of Math Destruction: How Big Data Increases Inequality and Threatens Democracy. Crown Publishing Group.
[33] Noble, S. U. (2018). Algorithms of Oppression: How Search Engines Reinforce Racism. NYU Press.
[34] Pasquale, F. (2015). The Black Box Society: The Secret Algorithms That Control Money and Information. Harvard University Press.
[35] Zuboff, S. (2019). The Age of Surveillance Capitalism: The Fight for a Human Future at the New Frontier of Power. PublicAffairs.

Como señala Martha Nussbaum, "la dignidad humana requiere que cada persona sea tratada como un fin en sí misma, no como un mero medio para los fines de otros"[36]. En la era de la IA, esto implica asegurar que los sistemas automatizados respeten y promuevan la dignidad humana, especialmente la de mujeres, en lugar de socavarla. Ello supone de forma directa poder asegurar en la era de la digitalización la igualdad y no discriminación para todos.

Helen Nissenbaum propone el concepto de "integridad contextual" como un marco para entender y proteger la privacidad en la era digital.[37] Esto es particularmente relevante en el contexto del uso de datos personales para entrenar y alimentar sistemas de IA. Especialmente en el uso de algoritmos en la toma de decisiones judiciales y administrativas, la predictibilidad plantea serios desafíos al debido proceso y a la posibilidad de impugnar decisiones injustas[38] .

Así pues, el peso de la dignidad humana ha de ser prevalente, representa un prius lógico, ontológico y deontológico que está justificado, no sólo desde la perspectiva histórica o filosófica, sino desde el reconocimiento jurídico que a nivel nacional, regional e internacional se ha dado, a través de las normas e instituciones que forman el Derecho nacional e internacional. Es decir, las normas e instituciones, que democráticamente hemos adoptado como limitación al poder y a la libertad humana, para lograr una convivencia pacífica en las comunidades donde hombres y mujeres desarrollamos nuestra existencia. Fuera de este contrato social, no hay “Derecho Justo” en el que podemos proyectar el bien común a las generaciones futuras. La tecnología, ni puede ni debe estar fuera de su alcance.

Admitir lo contrario sería reconocer el totalitarismo de un poder que se impone a la humanidad por encima de ella misma. Quizás la "pluma de los derechos" puede contrarrestar este riesgo.

[36]Nussbaum, M. C. (2006). Frontiers of Justice: Disability, Nationality, Species Membership. Harvard University Press.

[37] Nissenbaum, H. (2009). Privacy in Context: Technology, Policy, and the Integrity of Social Life. Stanford University Press.

[38] Citron, D. K., & Pasquale, F. (2014). The scored society: Due process for automated predictions. Washington Law Review, 89, 1.

IV. LA PLUMA DE LOS DERECHOS: EL DERECHOS HUMANOS DE IGUALDAD EN LA ERA DIGITAL

En el contexto de la Inteligencia Artificial y el sesgo algorítmico, los derechos de igualdad requieren una especial atención, que como explicaremos más adelante se ve fortalecida con nuevas interpretaciones de derechos que tienen sentido propio en la era digital. Nos encontramos con nuevas formas de discriminación que son difíciles de detectar y abordar con las leyes existentes[39], son más sutiles y generalizadas que la discriminación tradicional[40]. Esta posible violencia opaca y sutil que el sesgo algorítmico ejerce contra la igualdad llevan a Frank Pasquale afirmar que "No podemos proteger derechos que no podemos ver que están siendo violados"[41]. También resulta difícil concretar la responsabilidad, cuando se violentan debido a la complejidad de las cadenas de suministro y los sistemas de IA[42] , por lo que como indica Mireille Hildebrandt "nuestros sistemas legales y éticos deben volverse más ágiles para mantenerse al día con la innovación tecnológica" [43].

En cuanto a los derechos provenientes de la intimidad, la privacidad y la protección de datos de carácter personal, coincido con el concepto de "capitalismo de vigilancia" de Shoshana Zuboff,[44] por el que "los datos personales se han convertido en una nueva forma de capital. Resulta interesante al respecto la aportación de Helen Nissenbaum que propone la "integridad

[39] Wachter, S., Mittelstadt, B., & Russell, C. (2021). Why fairness cannot be automated: Bridging the gap between EU non-discrimination law and AI. Computer Law & Security Review, 41, 105567.
[40] Borgesius, F. Z. (2018). Discrimination, artificial intelligence, and algorithmic decision-making. Council of Europe, Directorate General of Democracy.
[41] Frank Pasquale afirma que "No podemos proteger derechos que no podemos ver que están siendo violados" Pasquale, F. (2015). The Black Box Society: The Secret Algorithms That Control Money and Information. Harvard University Press.
[42] Responsabilidad y rendición de cuentas: Kate Crawford y Vladan Joler, en su proyecto "Anatomy of an AI System", , destacando los desafíos para atribuir responsabilidad en casos de violaciones de derechos Crawford, K., & Joler, V. (2018). Anatomy of an AI System: The Amazon Echo As An Anatomical Map of Human Labor, Data and Planetary Resources. AI Now Institute and Share Lab.
[43] Mireille Hildebrandt “nuestros sistemas legales y éticos deben volverse más ágiles para mantenerse al día con la innovación tecnológica" Hildebrandt, M. (2015). Smart Technologies and the End(s) of Law: Novel Entanglements of Law and Technology. Edward Elgar Publishing.
[44] Zuboff, S. (2019). The Age of Surveillance Capitalism: The Fight for a Human Future at the New Frontier of Power. PublicAffairs.

contextual" por la que "la privacidad se entiende en términos de expectativas contextuales sobre el flujo de información, no simplemente en términos de control sobre los datos personales"[45]

Respecto a libertad de expresión en la era de las redes sociales y los algoritmos de moderación de contenido, Kate Klonick describe cómo se han convertido en "los nuevos gobernadores del discurso en línea", planteando preguntas sobre la libertad de expresión en espacios privados pero públicamente significativos en los que el control de la libertad de expresión y la diversidad de voces en línea pueden verse comprometidos.[46] Además, los algoritmos de recomendación en los medios no representan adecuadamente el contenido producido por mujeres, ya que normalmente no han tenido puestos de dirección en los medios de comunicación y el algoritmo tiene en cuenta este dato produciendo un sesgo.

Para reducir estos efectos se proponen diferentes soluciones de control para detectar y mitigar el sesgo en los sistemas de IA, especialmente destacamos las auditorías algorítmicas que pueden proporcionar una forma de responsabilidad algorítmica.

Además, resulta fundamental la exigencia de transparencia y comprensión de los sistemas de IA lo que supone un esfuerzo en educación y alfabetización digital que ha de dirigirse especialmente a las niñas y a las mujeres. Igualmente, la Gobernanza y el empoderamiento digital ciudadano fomenta la regulación proactiva, a la que se refiere Margrethe Vestager, Comisaria Europea de Competencia, argumentando que "necesitamos reglas que garanticen que la IA sea ética, segura y refleje nuestros valores democráticos", donde la igualdad de género tiene un protagonismo decisivo[47].

Veamos a continuación como los instrumentos de los Derechos humanos contemplan la igualdad, especialmente la de género, y como lo ha resuelto de forma magistral la Unión Europea mediante la Declaración Europea sobre los Derechos y Principios Digitales para la Década Digital, en la que expresa cómo han de aplicarse en el mundo digital los valores y derechos fundamentales que se aplican fuera de línea. Aspiran a promover una vía europea para la transformación digital, que sitúe a las personas en el centro, reafirme los derechos humanos universales y beneficie a todas las personas, mujeres y hombres, a las empresas y a la sociedad

[45] Nissenbaum, H. (2009). Privacy in Context: Technology, Policy, and the Integrity of Social Life. Stanford University Press.
[46] Klonick, K. (2018). The New Governors: The People, Rules, and Processes Governing Online Speech. Harvard Law Review, 131(6), 1598-1670.
[47] Vestager, M. (2019). Speech at the EU Conference on AI, Brussels, 18 October 2019.

en su conjunto, garantizando la contribución de la tecnología a la acción por el clima y la protección del medio ambiente.[48]

V.EL PAPIRO DE LAS LEYES: MARCO LEGAL Y DESAFÍOS

En el antiguo Egipto, el papiro era el soporte en el que se escribían las leyes y los textos sagrados. En nuestra analogía, el "papiro de las leyes" representa el marco de regulación actual y como las normas, la política y las instituciones abordan la discriminación de la IA a través del sesgo algorítmico.

Declaraciones e Instrumentos internacionales

La humanidad tuvo que luchar mucho hasta llegar, en 1949, a que el artículo 1° de la Declaración Universal de los Derechos Humanos de la ONU, reconociese que "Todos los seres humanos nacen libres e iguales en dignidad y derechos"[49].

Desde que se abolió la esclavitud hasta hoy en día, corren ríos de tinta en la justificación, exigencia y necesidad de garantizar el derecho a la igualdad y a la no discriminación, cuya primera definición en términos modernos e internacionales la encontramos en el Convenio formulado por la Organización Internacional del Trabajo en el año 1958, en el árticulo 1 de la Convenio 111, sobre la Discriminación. La definición allí formulada fue tomada posteriormente en la Convención relativa a la Lucha contra las Discriminaciones en la esfera de la Enseñanza, del año 1960 (art.1) y ampliada en el también árticulo primero de la Convención Internacional sobre la Eliminación de todas las Formas de Discriminación Racial. Se recogió con voz propia en el Pacto Internacional de Derechos Civiles y Políticos arts. 2 (1), 3, 20 (2); 23 (4), 24 (1), 26, y en el Pacto Internacional de Derechos Económicos, Sociales y Culturales artículos 2 (2), 3,10 (3). Con este mismo espíritu se recogió en la Organización Internacional del Trabajo (OIT)[50] y en la Declaración de Filadelfia de 1944. Más recientemente en el Objetivo 10 de

[48] Punto 5 de la declaración Aspiramos a promover una vía europea para la transformación digital basada en los valores europeos y los derechos fundamentales de la UE, que sitúe a las personas en el centro, reafirme los derechos humanos universales y beneficie a todas las personas, empresas y a la sociedad en su conjunto. el 23.1.2023 Diario Oficial de la Unión Europea C 23/1

[49] También los artículos 23 y 74 de la misma declaración, avalan el principio de igualdad

[50] Constitución de la Organización Internacional del Trabajo, Oficina Internacional del Trabajo, Ginebra 2010. http://www.ilo.org/public/spanish/bureau/leg/download/constitution.pdf

reducción de desigualdades, que pretende que las políticas económicas y sociales sean universales, prestando especial atención a las necesidades de las comunidades desfavorecidas y marginadas en la Agenda 20-30 de la Naciones Unidas.[51]

Los instrumentos regionales de Derechos humanos han expresado la igualdad y la no discriminación de las personas con especial protagonismo. En América, está recogida en el art 2 de la Declaración Americana de Derechos y Deberes del Hombre y en la Convención Americana de Derechos Humanos arts. 1 (1) (2), 13 (5), 17 (4) y 24. En la Carta Europea de Derechos Fundamentales se le dedica todo el capítulo III arts. del 20 al 26. La Carta africana sobre los derechos humanos y de los pueblos lo recoge en el art 2, también dentro del articulado de apertura. Respecto al sistema asiático de derechos humanos, se expresa en la Declaración de Derechos Humanos de la ASEAN de 2012, que lo recoge bajo la formulación del 1 y segundo principio general del texto[52]. Destacando la mención expresa de antidiscriminación a las minorías, mujeres, niños, trabajadores, personas con capacidades diferentes de la Carta Asiática de Derechos Humanos. Para terminar, no debemos olvidar la Carta Árabe de Derechos Humanos, que en su artículo 11 establece que todas las personas son iguales ante la ley y tienen derecho a disfrutar de su protección sin discriminación.

El Impacto del sesgo algorítmico en relación con la igualdad de género en relación con la igualdad de oportunidades, la igualdad racial, y de edad, además de las normas internas relativas a la no discriminación, violenta todo el corpus que como acabamos de comprobar, ha de tenerse en cuenta que ha sido adoptado por el Derecho interno de la casi totalidad de los Países de la Tierra.

Aquí me pregunto, si el mega sesgo colabora con la perpetuación y amplificación de las desigualdades, especialmente en lo que se refiere al mismo acceso en cuanto a la información solicitada. Pues como hemos visto en el debate con la IA el programa, juzga y toma decisiones propias al seleccionar la información que propone, lo que genera desconfianza[53] y representa un desafío para la democracia en el conocimiento y en el debate científico.

[51] Para una mayor información recomiendo el acceso a https://www.un.org/sustainabledevelopment/es/inequality/

[52] 1. Todas las personas nacen libres e iguales en dignidad y derechos. Están dotados de razón y conciencia y deben comportarse los unos con los otros con espíritu de humanidad.

[53] Danielle Keats Citron argumenta que la falta de transparencia y responsabilidad en los sistemas algorítmicos puede erosionar la confianza en las instituciones y en la tecnología misma Esto puede tener efectos profundos en la cohesión social y la participación democrática.

La propuesta de la UE para fortalecer la protección de los derechos humanos de igualdad en la era de la IA:

A pesar del impacto global de la regulación de la UE, a través del "Efecto Bruselas", y teniendo en cuenta los límites de este enfoque,[54] el Reglamento de Inteligencia Artificial (IA) [55] que ha entrado en vigor el día pasado 1 de Agosto, representa un ejemplo a seguir para las normas que tengan los derechos humanos como referentes para el desarrollo y despliegue de sistemas de IA.

Regula aspectos como el desarrollo, introducción en el mercado y utilización de los sistemas de IA, las principales prohibiciones de determinadas prácticas, los requisitos específicos para los sistemas de IA de alto riesgo y el resto de los sistemas de IA, también su control y la vigilancia, entendiendo que "los principios de derechos humanos deben guiar todo el ciclo de vida de los sistemas de IA"[56].

Constituye la primera regulación general a nivel mundial con jerarquía normativa de primer orden dentro del derecho de la Unión Europea, por lo que vincula directamente a todos los Estados miembros y con los que estos se relacionen. Los principales objetivos de esta regulación además de la protección de los derechos fundamentales descansan en la Transparencia y la rendición de cuentas. Así pues, a través de su articulado, el Reglamento garantiza que los sistemas de IA respeten los derechos individuales y no reproduzcan sesgos discriminatorios, por lo que nuestro sesgo de discriminación por razón de género quedaría minimizado.

Se establece la necesidad de información sobre su uso y propósito, así como los métodos empleados, lo que incluye la obligación de establecer marcos de rendición de cuentas para asegurar la responsabilidad de los desarrolladores y usuarios. Hay un especial tratamiento de la ciberseguridad y la legalidad con el compromiso de facilitar la inversión y la innovación en este campo.

[54] Bradford, A. (2020). The Brussels Effect: How the European Union Rules the World. Oxford University Press.
[55] El Diario Oficial de la Unión Europea (DOUE) publicó el pasado 12 de julio el Reglamento (UE) 2024/1689 del Parlamento Europeo y del Consejo, de 13 de junio de 2024, por el que se establecen normas armonizadas en materia de inteligencia artificial (Reglamento de Inteligencia Artificial. Esta norma del 1 de agosto de 2024 se aplicará un año después de su entrada en vigor, en agosto de 2025, y respecto a las obligaciones para los sistemas de alto riesgo en tres años. Estando supervisada por las autoridades nacionales, apoyadas por la oficina de IA dentro de la Comisión Europea.
[56] Council of Europe. (2019). Unboxing Artificial Intelligence: 10 steps to protect Human Rights.

Especialmente trata la discriminación cuando establece que quedan prohibidos los sistemas de inteligencia artificial de riesgo inaceptable.

Básicamente son los que tienen un elevado potencial de manipulación de personas a través de técnicas subliminares; También aquellos que aprovechan las vulnerabilidades de grupos específicos para alterar los comportamientos de las personas, pudiéndoles provocar perjuicios físicos o psicológicos; Igualmente los que permiten a las autoridades públicas evaluar o clasificar a las personas provocándoles un trato perjudicial o desfavorable tanto a ellas como a los colectivos a los que pertenecen; y finalmente los que posibilitan la identificación biométrica remota en tiempo real dentro de espacios públicos con la finalidad de aplicar la ley, salvo, en este último caso, que su uso sea necesario para alcanzar alguno de los objetivos de interés general, como la seguridad nacional prevista expresamente por la normativa europea.

En el caso de los sistemas de inteligencia artificial de alto riesgo, su comercialización está condicionada al cumplimiento de una serie de requisitos esenciales obligatorios, establecidos en el Reglamento y además que superen un examen técnico previo a dicha comercialización que demuestre su conformidad con dichos requisitos, lo que se consigue demostrando que los programas informáticos en cuestión han sido desarrollados siguiendo las exigencias técnicas establecidas, alternativamente, en normas armonizadas (fijadas por los organismos europeos de normalización), en especificaciones comunes (elaboradas por la Comisión) o a través de otras soluciones técnicas equivalentes (generadas muy a menudo por los propios operadores informáticos).

Finalmente, los sistemas de inteligencia artificial de bajo riesgo únicamente tienen la obligación de informar a los individuos de que están interactuando con un sistema de inteligencia artificial.

Para quienes incumplan la normativa se prevén multas con un rango que va desde los 35 millones de euros (37,6 millones de dólares) o el 7 % del volumen global de negocio de las empresas, hasta los 7,5 millones de euros (8 millones de dólares) o el 1,5 % del volumen global de negocio.

Algunos puntos débiles de la Ley podemos indicar que son la práctica de la autoevaluación de riesgos de las empresas de IA, ya que pudiera poner en peligro la protección de los derechos fundamentales, especialmente los de igualdad de género que aquí nos interesan. También la debilidad que representan los estándares para las evaluaciones de impacto sobre los

derechos fundamentales y la propia gobernanza pues a pesar de su regulación, no se han definido mecanismos concretos para garantizar la participación cívica en la implementación y aplicación de la Ley.

Para superar estos defectos se propone la prohibición absoluta del reconocimiento de emociones y limitar el uso biométrico, así como la creación de un registro público de algoritmos. Sería útil asegurar la igualdad en los sistemas de IA y fomentar la participación de la sociedad civil, incluyendo activamente aquellos como las mujeres, y las minorías, que pueden verse especialmente afectadas por estos sistemas. La cofundadora del European Laboratory for Learning and Intelligent Systems (ELLIS), Nuria Oliver, incide en que son varias las cuestiones que impactan sobre la falta de representatividad de la inteligencia artificial, destacando especialmente la falta de diversidad de género en la disciplina de la inteligencia artificial con apenas un 15 % de mujeres, en este sentido algunas voces indican que cerrar la brecha de género en las profesiones digitales llevará más de 30 años.

VI. LA RECONSTRUCCIÓN DEL EQUILIBRIO: LA VISIÓN DE LA MUJER Y LA JUSTICIA ALGORÍTMICA PARA TODOS.

Como dije en otra ocasión, derechos humanos, big data e inteligencia artificial son elementos de un complejo algoritmo jurídico, donde debemos acudir a los valores y principios inspiradores del Derecho para poder proponer soluciones. Libertad, Justicia e Igualdad deben estar aseguradas y concretadas en el trato digno a la persona en la era digital.

En la restauración del equilibrio respecto a la desigualdad de trato resulta fundamental como hemos visto anteriormente no solo que se oiga la voz de la mujer, lo que representa una posición pasiva, sino que las mujeres hablemos y que nuestra opinión y nuestras propuestas construyan pensamiento y definan soluciones a los problemas actuales.

La progresiva y definitiva incorporación de la mujer al pensamiento científico está siendo la mejor herramienta para reducir esa desigualdad irracional, castrante y ahora tecnológica.

La visión de la mujer en el mundo ayuda a corregir los defectos y errores de la sociedad actual, definida por valores líquidos y el individualismo hedonista. En este contexto, el

feminismo no sólo pretende corregir los sesgos machistas de los algoritmos, sino crear una tecnología inclusiva y representativa a través de la conocida como de justicia algorítmica.

La justicia algorítmica supone crear algoritmos capaces de mitigar los patrones de discriminación, existentes en los datos que se usan para procesar la información, es decir detectar los sesgos discriminatorios y contrarrestarlos computacionalmente.

En la actualidad ya hay herramientas que permiten al sistema tomar decisiones no sesgadas, de forma que se pueden extraerse variables de sexo, de raza, económicas etc, de manera que no los tenga en cuenta a la hora de tomar sus decisiones[57].

La IA puede ser una herramienta para la igualdad, usándola para corregir los sesgos humanos discriminatorios que hemos tenido hasta ahora. Puede contribuir a mejorar decisiones referenciadas por el machismo, incluso predecir situaciones injustas para poder cambiarlas.

Es preciso superar el sentido utilitarista de la tecnología, es necesario concretar para qué la queremos y tener en cuenta que una visión feminista de la tecnología, como la ecofeminismo lo hace en la ecología[58], ofrece una magnifica herramienta para cambiar radicalmente la vida de las personas, especialmente la de las mujeres.

En concreto la Justicia algorítmica supone, actuaciones de prevención, de acción y de sanción para corregir la discriminación del sesgo algorítmico, teniendo como referente el discurso de los derechos humanos y las reglas aplicables como el Reglamento de la Unión Europea sobre IA que promueve políticas corporativas que prioricen la equidad de género en sistemas de IA, exige transparencia, responsabilidad y establece sanciones.

[57] Un ejemplo basado en la propia experiencia personal de Marta Rodríguez-Tobalo, después de haber sido preguntada en varias ocasiones por su situación sentimental en varias entrevistas de trabajo ya llevo a fundar Blindstairs que utiliza la inteligencia artificial para combatir que sesgos discriminatorios como el sexo o el origen influyan a la hora de cribar candidaturas en los procesos de selección de personal. A través de la inteligencia artificial, se elimina de los currículum todos los datos personales para evitar que se pueda hacer una trazabilidad de la persona candidata y se pasa todo a lenguaje neutro.Así, la persona de recursos humanos encargada del proceso de selección no sabe si es un hombre o una mujer. Una vez se llega a la fase de las entrevistas, el sistema crea un avatar de rasgos neutros que, gracias a tecnología de reconocimiento facial y de modulación de la voz, evita que se pueda diferenciar el género a la vez que replica gestos y expresiones de la persona candidata. Otro caso práctico en el que la tecnología contribuye a la diversidad y a corregir la desigualdad es el de Themis, un software de procesamiento de lenguaje natural que revisa cualquier texto para convertirlo en inclusivo: detecta expresiones no inclusivas y sugiere alternativas para cambiarlas.

[58] Véase Sustainable Prosperity and Circular Economy in the Care of Mother Earth — The Blue Kiss en https://www.qeios.com/read/MMRNCV

Hacer más transparentes los procesos de toma de decisiones algorítmicas para usuarios y partes interesadas, reduciendo el sesgo discriminatorio del género supone en primer lugar promover la inclusión de mujeres y otros grupos subrepresentados en los equipos de desarrollo de IA, así como fomentar una cultura de conciencia sobre cuestiones de género en los ambientes de trabajo a través de programas de capacitación sobre sesgo algorítmico y ética de la IA para desarrolladores y directivos implicados en la toma de decisión.

En segundo lugar, respecto a la programación en sí misma, se precisan actuaciones previas de revisión y corrección de datos para ajustar los algoritmos. Son necesarios análisis previos de los conjuntos de datos de entrenamiento que identifiquen y eliminen sesgos de género existentes. Además, establecer mecanismos que garanticen que los datos sean representativos e incluyan una diversidad adecuada de géneros.

Se han de implementar técnicas de mitigación de sesgos durante el proceso de entrenamiento del modelo y utilizar algoritmos de "debiasing" para reducir sesgos aprendidos, independientemente de la calidad de los datos.

La Justicia algorítmica para ser efectiva necesita una actividad de policía que se consigue a través de un monitoreo continuo para detectar y corregir sesgos emergentes durante la operación de los sistemas de IA, ello supone realizar auditorías regulares para evaluar los resultados de acuerdo con métricas de equidad.

Es importante notar, que este es un proceso continuo que requiere vigilancia constante y adaptación a medida que se identifican nuevas formas de sesgo y las tecnologías evolucionan. De hecho, como indica la Exposición de Motivos del Reglamento Europeo de inteligencia artificial, la definición jurídica de la inteligencia artificial, lejos de estar cerrada, se ajusta en función de la evolución tecnológica.

Para su desarrollo sería deseable una legislación específica sobre el sesgo algorítmico y la discriminación de género en sistemas de IA, que estableciese requisitos legales para reducir al mínimo el riesgo de discriminación algorítmica, especialmente en el diseño y calidad de los conjuntos de datos. También actualizar las leyes existentes de no discriminación para incluir explícitamente la discriminación algorítmica basada en género e implementar procedimientos propios para impugnar decisiones algorítmicas discriminatorias.

Para finalizer debemos considerar nuevos derechos, como el derecho a la explicación algorítmica o el derecho a la interacción humana significativa como indica Risse, profesor de

Filosofía y Políticas Públicas en Harvard, que podrían considerarse como parte de los nuevos derechos fundamentales, ya no de los hombres y mujeres, sino como en el caso del medio ambiente, derechos de la humanidad presente y futura.

Muchas gracias.

FUENTES

— Ajunwa, I. (2020). The Paradox of Automation as Anti-Bias Intervention. Cardozo Law Review, 41(5), 1671-1742.

— Angwin, J., Larson, J., Mattu, S., & Kirchner, L. (2016). Machine Bias. ProPublica.

— Bogen, M., & Rieke, A. (2018). Help Wanted: An Examination of Hiring Algorithms, Equity, and Bias. Upturn.

— Borgesius, F. Z. (2018). Discrimination, artificial intelligence, and algorithmic decision-making. Council of Europe, Directorate General of Democracy.

— Bradford, A. (2020). The Brussels Effect: How the European Union Rules the World. Oxford University Press.

— Buolamwini, J., & Gebru, T. (2018). Gender Shades: Intersectional Accuracy Disparities in Commercial Gender Classification. Proceedings of the 1st Conference on Fairness, Accountability and Transparency, 77-91.

— Citron, D. K., & Pasquale, F. (2014). The scored society: Due process for automated predictions. Washington Law Review, 89, 1.

— Council of Europe. (2019). Unboxing Artificial Intelligence: 10 steps to protect Human Rights.

— Crawford, K. (2021). Atlas of AI: Power, Politics, and the Planetary Costs of Artificial Intelligence. Yale University Press.

— Crawford, K., & Joler, V. (2018). Anatomy of an AI System: The Amazon Echo As An Anatomical Map of Human Labor, Data and Planetary Resources. AI Now Institute and Share Lab.

— De Andrés, T. (2002). Homo Cybersapiens. La inteligencia artificial y la humana. Pamplona, España, EUNSA, 90.

— Eubanks, V. (2018). Automating Inequality: How High-Tech Tools Profile, Police, and Punish the Poor. St. Martin's Press.

— Ferrajoli, L. (2006). Derechos y Garantías. La ley del más débil. Madrid: Editorial Trotta , págs. 78-79.

— Habermas, J. (2010). The concept of human dignity and the realistic utopia of human rights. Metaphilosophy, 41(4), 464-480.

— Hildebrandt, M. (2015). Smart Technologies and the End(s) of Law: Novel Entanglements of Law and Technology. Edward Elgar Publishing.

— Hurley, M., & Adebayo, J. (2016). Credit Scoring in the Era of Big Data. Yale Journal of Law and Technology, 18(1), 148-216.

— Keyes, O. (2018). The Misgendering Machines: Trans/HCI Implications of Automatic Gender Recognition. Proceedings of the ACM on Human-Computer Interaction, 2(CSCW), 1-22.

— Klonick, K. (2018). The New Governors: The People, Rules, and Processes Governing Online Speech. Harvard Law Review, 131(6), 1598-1670.

— Lessig, L. (1999). Code and Other Laws of Cyberspace. Basic Books.

— Mendez Rocasolano M. (2024). Sustainable Prosperity and Circular Economy in the Care of Mother Earth — The Blue Kiss acceso en https://www.qeios.com/read/MMRNCV

— Mendez Rocasolano M. (2022). Human Rights, Big Data and Artificial Intelligence: Elements of a Complex Algorithm. Security and Defence: Ethical and Legal Challenges in the Face of Current Conflicts, 93–102.

— Narayanan, A. (2013). Online Price Discrimination: Conspicuous by its Absence. 33 Bits of Entropy.

— Nissenbaum, H. (2009). Privacy in Context: Technology, Policy, and the Integrity of Social Life. Stanford University Press.

— Noble, S. U. (2018). Algorithms of Oppression: How Search Engines Reinforce Racism. NYU Press.

— Nussbaum, M. C. (2006). Frontiers of Justice: Disability, Nationality, Species Membership. Harvard University Press.

— Nussbaum, M. C. (2011). Creating capabilities: The human development approach. Harvard University Press.

— O'Neil, C. (2016). Weapons of Math Destruction: How Big Data Increases Inequality and Threatens Democracy. Crown Publishing Group.

— Pasquale, F. (2015). The Black Box Society: The Secret Algorithms That Control Money and Information. Harvard University Press.

— Perz Luño, A.(1987). "Sobre la igualdad en la Constitución española", en el Anuario de Filosofía del Derecho , vol IV, 133 y ss.

— Reglamento (UE) 2024/1689 del Parlamento Europeo y del Consejo, de 13 de junio de 2024, por el que se establecen normas armonizadas en materia de inteligencia artificial (Reglamento de Inteligencia Artificial.

— Rice, L., & Swesnik, D. (2013). Discriminatory Effects of Credit Scoring on Communities of Color. Suffolk University Law Review, 46, 935-966.

— Rosenblat, A. (2018). Uberland: How Algorithms Are Rewriting the Rules of Work. University of California Press.

— Russell, S. (2019). Human Compatible: Artificial Intelligence and the Problem of Control. Viking.

— Stucke, M. E., & Ezrachi, A. (2017). How Digital Assistants Can Harm Our Economy, Privacy, and Democracy. Berkeley Technology Law Journal, 32, 1239-1299.

— Velázquez Fernánde,H. (2019). ¿Qué tan natural es la inteligencia artificial? Sobre los límites y alcances de la biomímesis computacional, Colegio de Bioética de Nuevo León, México Naturaleza y Libertad. Número 12, 256.

— Vestager, M. (2019). Speech at the EU Conference on AI, Brussels, 18 October 2019.

— Wachter, S., Mittelstadt, B., & Russell, C. (2021). Why fairness cannot be automated: Bridging the gap between EU non-discrimination law and AI. Computer Law & Security Review, 41, 105567.

— Zuboff, S. (2019). The Age of Surveillance Capitalism: The Fight for a Human Future at the New Frontier of Power. PublicAffairs.