

Discriminación algorítmica e ilusoria neutralidad. Actuales desafíos ante el fortalecimiento de tradicionales desigualdades en materia de violencia contra la mujer

Algorithmic discrimination and illusory neutrality.
Current challenges in the face of reinforcing traditional
inequalities in violence against women

SANDRA LÓPEZ DE ZUBIRÍA DÍAZ

*Profesora contratada doctor interina del Departamento
Derecho público II y Filología I (área Derecho penal).
Miembro del Grupo de investigación de alto rendimiento
en libertad, seguridad y ciudadanía en el orden
internacional de la Universidad Rey Juan Carlos
ORCID: 0000-0002-5833-4375*

Recibido: 15/1/2024

Aceptado: 2/5/2024

doi: 10.20318/femeris.2024.8641

Resumen. El empleo de herramientas de inteligencia artificial aplicadas a diversos ámbitos constituye una realidad patente, conforme su uso crece exponencialmente con carácter general en múltiples ámbitos de la vida. Al respecto, se presenta aquí un análisis sobre aquellas cuestiones más alarmantes en materia preventiva e intervencionista, con especial vinculación al ámbito de la violencia de género. Entre otras, las que derivan de encontrarnos ante un sistema significativamente opaco, pero sobre el que se confía cierta neutralidad que puede convertirse en un instrumento capaz de obstaculizar la igualdad, a través de la proliferación de tradicionales discriminaciones enmascaradas en ilusorias neutralidades promovidas por herramientas generadas a partir de una base sesgada y excluyente. Con ello, lejos de combatir este fenómeno delictivo, estaremos fomentando sus bases patriarcales desigualitarias a través de la creación de herramientas tecnológicas teóricamente neutrales pero discriminatorias en la práctica. Al respecto, sin ánimo de desarrollar una postura negacionista ante la evolución tecnológica en este texto, lo que se realiza aquí es un estudio sobre aquellos aspectos que deben atenderse para que dicho avance evite el crecimiento de discriminaciones y violencias simbólicas, base de la violencia contra las mujeres, enmascaradas en una suerte de inteligencia artificial, neutral, que realmente esconda marcados sesgos de género.

Palabras clave: Algoritmo, inteligencia artificial, violencia, mujer, género.

*sandra.lopezdezubiria.diaz@urjc.es

Abstract. The use of artificial intelligence tools applied to various fields is a reality, as their use grows exponentially in many areas of life. In this respect, we present here an analysis of the most alarming issues in terms of prevention and intervention, with special reference to the field of gender violence. These include, among others, those deriving from the fact that we are faced with a system that is significantly opaque, but which is trusted with a certain neutrality that can become an instrument capable of hindering equality, through the proliferation of traditional discriminations masked in illusory neutrality promoted by tools generated from a biased and excluding base. In this way, far from combating this criminal phenomenon, we will be promoting its unequal patriarchal bases through the creation of technological tools that are theoretically neutral but discriminatory in practice. In this respect, without wishing to develop a denialist stance towards technological evolution in this text, what is done here is a study of those aspects that must be addressed so that such progress avoids the growth of discrimination and symbolic violence, the basis of violence against women, masked in a kind of artificial intelligence, neutral, that really hides marked gender biases.

Keywords: Algorithm, artificial intelligence, violence, women, gender.

1. Apunte inicial sobre la inteligencia artificial: de la ausencia de conceptualización unívoca a la proliferación de sus usos

Pese a que se puedan identificar referencias anteriores a ese momento, lo cierto es que es en 1955, en la Conferencia de Dartmouth (Parsheera, 2018), cuando se entiende adoptado terminológicamente el concepto de “inteligencia artificial”, a manos del profesor John McCarthy.

Desde entonces, hasta la actualidad, pese a la significativa evolución, especialmente en épocas recientes, la conceptualización no se ha asentado, destacando la ausencia de unanimidad en torno a este término (Hernández Giménez, 2019) y la proliferación de muy diversas definiciones sobre el término de IA (como examina Valls Prieto, 2021) que, si bien supone una riqueza conceptual, dificulta su aproximación, así como su regulación.

No obstante, el estudio sobre esta sí se ha potenciado, por lo que, pese a la inexistencia de consenso en cuanto a la terminología, sí que se han realizado numerosos estudios en cuanto a su contenido, donde se acuerda que nos encontramos ante “sistemas que manifiestan un comportamiento inteligente” (Borges Blázquez, 2020, p. 55).

En este sentido, pese a diversas categorizaciones existentes, destaca la distinción entre dos tipos de inteligencia artificial. Por un lado, la del tipo débil (*o weak AI*) –que sería equiparada con las capacidades humanas– por otro, la de tipo fuerte (*o strong AI*) –que superaría aquellas–. Sin embargo, y pese a las posibilidades, en la actualidad únicamente se ha desarrollado la del primer tipo (Miró Llinares, 2020, p. 84), a través de *machine learning*, *Data mining*, *Procesamiento del Lenguaje Natural*, entre otros “campos de investigación que se dedican a explorar la automatización de los procesos del intelecto humano” (Delgado Calvo, 1996, p. 17)

Como se señalaba, pese a la diferenciación entre ambos tipos, hoy en día el desarrollo de esta inteligencia se ha visto limitado a la de tipo débil, constituyéndose especialmente a través de la expansión de algoritmos de predicción que potencian la obtención de un resultado determinado con base en operaciones matemáticas entrelazadas, a través del uso de una cantidad ingente de datos.

Con ello, doctrinalmente se circunscribe el entendimiento de la inteligencia artificial a una “combinación de algoritmos planteados con el propósito de crear máquinas que presenten las mismas capacidades que el ser humano” (Muñoz Rogríguez, 2020, p. 699) que han sido creados a través de la incorporación de grandes cantidades de datos que “ordenados de manera comprensible (Smart data), un modelo matemático va utilizando de manera aleatoria, hasta establecer patrones de correlación determinista entre ellos” (De Miguel Beriain, 2019, p. 536).

No obstante, pese a que parezca que el desarrollo actual es simplista, los retos que plantea hoy en día son numerosos¹, lo cual permite deducir que, en un futuro, con una inteligencia artificial más desarrollada, los desafíos serán notables. Entre otros, MIRÓ LLINARES (2018) destaca los cuestionamientos frente a la exigencia de responsabilidad penal de objetos automatizados o probable vulneración de garantías procesales, entre otros aspectos relevantes en la materia que pueden afectar especialmente al ámbito del Derecho penal y procesal, por cuanto “las formas en las que la inteligencia artificial puede repercutir en la comisión de hechos delictivos resultan muy plurales y diversas” (Morillas Fernández, 2023, p. 71) lo que está llevando a una nueva categoría terminológica en torno a los *AI-Crime* e incluso a la revisión de la teoría jurídica del delito para valorar el encaje de la responsabilidad penal de las personas jurídicas ante los nuevos desafíos que plantean los daños punibles que son consecuente del uso de la IA (Del Rosal Blasco, 2023).

De los numerosos riesgos advertidos surge la necesidad de una adecuada regulación, como ya se plantea desde el Parlamento Europeo y el Consejo, dada su voluntad de crear un reglamento en materia de inteligencia artificial que aglutine normativa y que prevenga las diversas amenazas observadas. En esta línea, pese a la inexistencia todavía de un reglamento marco en la materia, España es ya pionera, tras su Real Decreto 729/2023, de 22 de agosto, por el que se aprueba el Estatuto de la Agencia Española de Supervisión de Inteligencia Artificial, al crear el primer organismo en el marco de la UE enfocado en la supervisión de la IA. Así, entre otros, se establecen

tareas de supervisión, el asesoramiento, la concienciación y la formación dirigidas a entidades de derecho público y privado para la adecuada implementación de toda la normativa nacional y europea en torno al adecuado uso y desarrollo de los sistemas de inteligencia artificial, más concretamente, de los algoritmos. Además, la Agencia tendrá la función de inspección, comprobación, sanción y demás que le atribuya la normativa europea que le resulte de aplicación y, en especial, en materia de inteligencia artificial (art. 4 Real Decreto 729/2023, de 22 de agosto)

Y es que los algoritmos se encuentran presentes en muy diversos ámbitos de la vida. Entre otros, el contexto de la comunicación es aquel que penetra especialmente en la población en general, por cuanto el uso de plataformas y redes sociales es continuo (como

¹ No solo aquellos destacados por juristas, como se expondrá en este texto, sino los que los propios expertos en su desarrollo han advertido, sorprendiendo las peticiones emanadas en pro de un tiempo de espera en el avance de la IA, dados los peligros advertidos. Sobre esta cuestión en la prensa: <https://elpais.com/tecnologia/2023-03-29/expertos-en-inteligencia-artificial-reclaman-frenar-seis-meses-la-carrera-sin-control-de-los-chatgpt.html>

puede ser twitter, Instagram o Facebook; e incluso aquellas desarrolladas para el emparejamiento, como pueden ser Tinder, Badoo o grinder).

En este mismo campo incluso los programas de televisión están apostando por la aplicación de inteligencia artificial –como es el caso del programa “falso amor”– que, a través de tecnología *deepfake*, pone en tela de juicio el comportamiento de los concursantes ante vídeos que pueden no haber sido protagonizado por estos, sino creados a través de esta tecnología². Por otro lado, nos encontramos ante el desarrollo de numerosos chatbot como asistentes virtuales en prácticamente cualquier entidad, pública o privada que tienden a sustituir la atención personalizada por parte de agentes para la resolución de consultas sencillas.

Con todo, nos encontramos con el uso de plataformas a través de las cuales se naturaliza la atribución de decisiones tan relevantes como las amorosas, condicionando las personas con las que poder “hacer match” en función de la decisión tomada por un algoritmo, con un funcionamiento significativamente opaco.

Asimismo, con el uso de las redes sociales, la magnitud de datos ofrecidos se multiplica, devolviendo “casuales” recomendaciones de páginas o productos a los usuarios, gracias a los datos ofrecidos “libremente” por estos, lo cual no influye directamente en el ámbito económico y el consumo, sino que se han advertido incluso perniciosas consecuencias sociales, como son la polarización social y el incremento del extremismo (Rodríguez-Gordo y Sarrión-Andaluz, 2021, p. 166) e incluso el ámbito de las series/ programas de televisión han encontrado un campo en desarrollo acudiendo al uso de inteligencia artificial para “modernizar” tradicionales formatos que no solo potencian relaciones tóxicas, sino que potencian la creatividad de los delincuentes al encontrar nuevas formas para la suplantación de identidad, entre otros tipos delictivos, que pueden ser más fácilmente ejecutados con este tipo de herramientas que, además, resultan de mayor dificultad para su persecución, mayor anonimización para sus autores y, mayor perjuicio para sus víctimas.

Con todo, la aplicación de este tipo de herramientas no se limita al ámbito de las redes sociales, las diversas plataformas de emparejamiento o los medios de comunicación en general, sino que invaden otros contextos de especial relevancia, como el económico, el de la salud, la administración pública y, en concreto, el ámbito judicial.

Así, la inteligencia artificial nutre el ámbito económico para acciones tan diversas como el asesoramiento financiero, la valoración de riesgo o el control de blanqueo de capitales, entre otros (Fernández, 2019) y parece pisar fuerte en el campo de la salud, donde se alude a la posibilidad de reducir significativamente el coste de atención, optimizar la investigación y mejorar los resultados en salud de los pacientes (Sánchez Rosado y Díez Parra, 2022, p. 130 y ss.)

² Donde además de potenciar aspectos negativos vinculados con las relaciones tóxicas y el mito del amor romántico –como los celos, la inseguridad, la desconfianza o la posesión– se exponen las posibilidades que ofrece la inteligencia artificial para la suplantación de la identidad de las personas, lo que potencia el desarrollo de nuevas criminalidades, frente a lo que el derecho penal y el derecho procesal, también tendrá que adaptarse. Entre otros ejemplos que ya son realidades a las que atender: <https://elpais.com/espana/2023-09-18/la-policia-investiga-el-desnudo-integral-de-varias-menores-en-extremadura-con-inteligencia-artificial-me-dio-un-vuelco-el-corazon.html>

Asimismo, las posibilidades en el campo judicial parecen ser numerosas, si bien su desarrollo actual se centra en cuestiones como el análisis de datos, el estudio de la jurisprudencia o incluso las “técnicas de justicia predictiva” de las compañías aseguradoras para valorar el coste de oportunidad de acudir a juicio (Simón Castellano, 2021, p. 6). No obstante, el avance en esta materia está sobre la mesa, por cuanto el Ministerio de Justicia ha potenciado recientemente su uso, esencialmente a través de herramientas de ayuda a la anonimización de documentos y autocatalogación de estos a través del recurso a la IA con el objetivo de mejorar los servicios de manera más eficiente³, pero con el convencimiento de que las funciones se irán incrementando a medida que se asiente su uso en estos ámbitos.

Con relación a los contextos señalados, es interesante plantear una mirada general a otros países, como China, EEUU o Reino Unido, donde su uso es más significativo y donde este se ha traducido rápidamente en numerosos cuestionamientos éticos y legales. Como ejemplos a destacar, se encuentra COMPAS –siendo una herramienta de predicción de reincidencia cuyos resultados se obtienen a partir de datos relativos a sentencias ejecutadas– o FAST –constituyéndose como mecanismo preventivo de atentados terroristas en el medio aéreo–. Asimismo, se encuentra CLOUD WALK como herramienta de predicción delictiva con carácter general, a través del uso de reconocimiento facial y observación de la conducta, que permite analizar la posible predisposición delictiva de la ciudadanía.

De la misma manera, ya en el ámbito penitenciario, también se encuentran STATIC 99 o HCR-20, como predictores de acciones violentas en este contexto. Asimismo, ya en un contexto ajeno a la predicción de delincuencia y/o reincidencia, también se aprecian sistemas para el “control” del comportamiento de la ciudadanía, como “Citizen Score”. Este sistema, aplicado en China, supone una suerte de marcador cuya puntuación aumenta o disminuye, entre otros ejemplos, en función de las actividades que realicen o de las personas de las que se rodeen, siendo premiados o castigados atendiendo a las puntuaciones obtenidas.

Como se adelantaba, estas herramientas han tenido mayor acogida en otros países, no siendo desarrolladas –no, al menos, de la misma manera, ni con el mismo impulso– en el contexto español. No obstante, sí encontramos algunos ejemplos con cierta similitud, como puede ser la herramienta RisCanvi, elaborada para hacer prevenciones de reincidencia en el seno de los centros penitenciarios de Cataluña, siendo presumible que, dadas las posibilidades que ofrece la inteligencia artificial, su uso sea multiplicado en los próximos tiempos.

Con todo, de este apunte inicial es posible inferir que en el ámbito de la inteligencia artificial se cuenta con mayor desconocimiento, que certezas, por cuanto ni siquiera el propio concepto resulta concreto. No obstante, esto no impide la proliferación de su uso en la actualidad, con la implementación de sistemas algorítmicos con capacidad de aprendizajes y de actuación automática que se postulan como auxilio en la toma de decisiones de muy diversos contextos, cuando no como sustitución de los profesionales.

Entre otros ámbitos, es posible apreciar el de la prevención y tratamiento de la violencia contra la mujer, donde la implementación de algoritmos ya ha tenido una aplicación

³ <https://www.mjusticia.gob.es/es/institucional/gabinete-comunicacion/noticias-ministerio/ia>

temprana y donde se aprecia un crecimiento en el interés del recurso a este tipo de herramientas, con numerosas propuestas en marcha, pese a sus cuestionamientos éticos y posibles problemas prácticos, que se atenderán en las siguientes líneas.

2. Algoritmización de la violencia contra la mujer: pasado, presente y futuro

Cuando se hace referencia al término de violencia contra la mujer, pese a la disparidad de terminología aplicada en este campo que, en ocasiones, puede desvirtuar el contenido en sí del concepto⁴, hacemos referencia a lo señalado en su art. 3 por el Convenio del Consejo de Europa sobre prevención y lucha contra la violencia contra las mujeres y la violencia doméstica señalando que

por “violencia contra las mujeres” se deberá entender una violación de los derechos humanos y una forma de discriminación contra las mujeres, y designará todos los actos de violencia basados en el género que implican o pueden implicar para las mujeres daños o sufrimientos de naturaleza física, sexual, psicológica o económica, incluidas las amenazas de realizar dichos actos, la coacción o la privación arbitraria de libertad, en la vida pública o privada.

A la vez que completa su definición añadiendo que

por “violencia contra las mujeres por razones de género” se entenderá toda violencia contra una mujer porque es una mujer o que afecte a las mujeres de manera desproporcionada.

Lo que, en definitiva, alude a una manifestación de violencia que supone una clara vulneración de los derechos humanos que encuentra muy diversos actos en los que concretarse (sean de carácter físico, sexual, económico, etc.) con particularidades propias derivadas de su perpetración de manera desproporcionada contra las mujeres donde es, precisamente el género –como roles y estereotipos asociados a la mujer– lo que condiciona la violencia.

En este orden de cosas, una vez que se ha hecho una sucinta aproximación conceptual sobre la materia, conviene atender a su vinculación con el objeto de estudio, como es la inteligencia artificial y su implicación en el campo de la violencia contra las mujeres.

Pese a que el interés por la inteligencia artificial y la proliferación de herramientas basadas en esta parece reciente, lo cierto es que su uso acompaña el ámbito de la violencia contra la mujer desde hace más de una década, encontrando que la algoritmización en este campo tiene un significativo recorrido.

En concreto, a través de un sistema de seguimiento integral en los casos de violencia de género, conocido como VIOGEN, cuyo origen data de 2007 y que se constituye como

⁴ Conviviendo tanto en su comprensión social, como en cuestión normativa, diversas terminologías como puede ser la violencia machista, la violencia patriarcal, la violencia de género –término asentado internacional y nacionalmente– la violencia por razones de género, entre otras. Bien es cierto que estas terminologías responden a miradas poliédricas que, en ocasiones, son utilizadas para ofrecer mayor peso a las referencias culturales del machismo o a las relaciones jerarquizadas de género que fundamentan el sistema patriarcal, sin embargo, se considera que la coexistencia de diversas nomenclaturas desconciertan la comprensión del fenómeno en su totalidad.

una aplicación informática que cuenta con el empleo de inteligencia artificial. Este sistema actúa cual red seguimiento, englobando a las diversas instituciones públicas con competencia en este campo (como la Policía Nacional, la Guardia Civil, las policías autonómicas, Instituciones penitenciarias, el Ministerio Fiscal o el Poder judicial), permitiendo una visión global de cada caso y planteando una protección más completa.

En el seno de este sistema se encontrarían una valoración policial de riesgo (VPR), así como una valoración policial de la evolución de dicho riesgo (VPER), siendo herramientas complementarias donde si bien en un primer momento se valora el nivel de riesgo apreciado, posteriormente se produce una valoración para advertir si se producen nuevas situaciones que puedan afectar al riesgo inicialmente observado y requiriera algún tipo de modificación en cuanto al protocolo de protección y prevención a seguir.

Ambos sistemas se fundamentan, esencialmente, en un algoritmo que, según la información introducida por los agentes –tras cumplimentar una serie de ítems basadas en las respuestas ofrecidas por la víctima– generan una respuesta en torno al riesgo que es estimado, dados los datos introducidos. Así, la víctima puede estar ante un riesgo “no apreciado”, “bajo”, “medio”, “alto” y “extremo” que condicionarán las medidas de protección asignadas a la víctima, desde una mera información de derechos o recursos de ayuda, hasta una vigilancia constante, de darse riesgo extremo.

Pese al dilatado recorrido de esta herramienta, sobre la cual se han realizado diversas mejoras en los últimos años, lo cierto es que sigue planteando problemas que deben ser atendidos.

Entre otros, el alto porcentaje de riesgo “no apreciado” –en torno a un 45% de los casos denunciados– lo cual no parece casar con la magnitud del problema, más cuando se han dado diversas situaciones en las que víctimas que habían sido identificadas como de riesgo “no apreciado” o “bajo”, finalmente fueron asesinadas.

En esta línea, parece que quizás uno de los problemas venga derivado de los datos introducidos en el sistema. Así, las mujeres que acuden a denunciar, generalmente lo hacen tras una situación de violencia, encontrándose en un estado que puede no resultar óptimo para dar respuestas idóneas –especialmente cuando los ítems requieren respuestas binarias–. Por ello, la formación de los profesionales es esencial, por cuanto deben traducir relatos de la víctima en respuestas dicotómicas de sí/no, en muchos casos de difícil interpretación, dado el estado de shock en el que puedan encontrarse si, efectivamente, la denuncia se produce inmediatamente después de una agresión.

En esta línea, se propone que la víctima cuente con adecuada asistencia legal y psicológica (Fundación Ana Bella, 2022, p. 33) desde el momento en el que decide interponer la denuncia, para tener una información adecuada de sus derechos y, especialmente, para contar con la ayuda necesaria para que pueda proporcionar un relato acomodado a la situación de violencia sufrida. Es cierto que actualmente las víctimas de este tipo de violencia cuentan con la posibilidad de acceder a la justicia gratuita, recibiendo asistencia inmediata de un profesional de la abogacía. No obstante, el recurso al acompañamiento psicológico no se encuentra debidamente atendido, lo cual resulta esencial para que la víctima pueda enfrentarse a la denuncia de la forma más adecuada, así como de continuar con el proceso.

Asimismo, indiscutiblemente, dada la magnitud de casos estudiados⁵ y la vigencia de este sistema, los datos con los que trabaja son muy numerosos. Sin embargo, cuando las experiencias no responden a escenarios previamente estudiados, el resultado ofrecido por el algoritmo puede no ser adecuado.

En este sentido, parece que no hay ejemplo más claro que el relativo a la crisis sanitaria producida por el covid, donde una situación de confinamiento total, miedo ante la incertidumbre, crisis económica, pérdida de empleo, desconocimiento sobre el virus, etc. suponen factores fundamentales que condicionan una determinada situación de riesgo que, al no tener datos previos de situaciones similares, no son tenidos en cuenta por el algoritmo, pese a que su relevancia sea evidente.

De la misma manera, se ha observado que, pese a la posibilidad ofrecida al profesional de modificar el riesgo apreciado (a un riesgo superior), en escasas ocasiones se producen variaciones frente a los resultados ofrecidos por el sistema, por lo que siempre es superior al 94% los casos en los que los agentes se conforman con el resultado ofrecido por la aplicación (Zurita, 2014, p. 777).

De este modo, nos encontramos ante una artificiosidad de la herramienta, sustituyendo la decisión del agente dada una errónea expectativa sobre la predicción que obstaculiza la toma de decisiones del agente, confiando obnubiladamente en el resultado ofrecido.

Con todo, la herramienta VIOGEN presenta grandes ventajas y, además, está sujeta a revisiones. Sin embargo, estas cuestiones deben ser tenidas en cuenta para que los datos introducidos sean lo más adecuados posibles –cuidando la asistencia a la víctima para que su declaración sea lo más fructífera posible y, siempre que sea posible, no sea la única, sino que se acompañe de otros testimonios– así como para reforzar la profesionalización de los agentes en este ámbito –que evite la sustitución de la opinión profesional por la decisión de un algoritmo, siendo este un respaldo y no un reemplazo–.

Pese a que el mayor recorrido en este ámbito lo ocupe el sistema VIOGEN, lo cierto es que en el presente se están desarrollando diversas herramientas a través del uso de inteligencia artificial que conviene señalar.

Entre otras, destaca el proyecto “Certeza de voz” en el que colabora el Instituto Andaluz de la Mujer, el cual está investigando en la detección precoz de víctimas a través del análisis de patrones de voz. De esta forma, se busca que, en las llamadas a servicios sanitarios andaluces, por diversas patologías, pueda filtrarse la posible existencia de casos de violencia contra la mujer, enmascarados en estas patologías, a través del discurso ofrecido por parte de las mujeres al expresar cuáles son los síntomas, cómo se han causado, qué palabras empleen o los silencios en su comunicación, entre otros.

De esta manera, se posibilitaría una atención temprana a posibles víctimas de violencia que ni siquiera hayan advertido su condición de víctimas o que, habiéndolo advertido, no consultan con los servicios de salud por estos motivos. Así, se permitiría no solo el auxilio a las mujeres para una adecuada identificación del contexto violento en el que

⁵ Actualmente, superando los 6 millones de valoraciones de riesgo. <https://www.lamoncloa.gob.es/serviciosdeprensa/notasprensa/interior/Paginas/2023/190523-sistema-viogen-victimas-violenciagenero.aspx#:~:text=En%20la%20actualidad%20el%20sistema,y%2033.548%20en%20riesgo%20bajo.>

se encuentran –desde las primeras fases– sino la reducción de diagnósticos incorrectos sobre patologías que, en realidad, esconden un origen de maltrato y que, sin conocerlo, pueden enmascarse en patologías “comunes” de ansiedad, depresión o estrés.

En términos similares, nos encontramos ante el proyecto empatía, el cual ha desarrollado “bindi”, siendo este un algoritmo que busca desarrollar indicadores de alerta ante contextos de riesgo. En este caso, en lugar de un análisis conversacional, se plantea la monitorización de emociones de víctimas que permitan advertir patrones de comportamiento.

De este modo, analizando cuáles son las emociones desarrolladas, se pretende establecer un mecanismo preventivo que advierta una posible situación de riesgo en la que esta se encuentre, avisando a terceros –como a servicios de emergencia o a familiares– posibilitando medidas de seguridad ante una situación de amenaza. Por tanto, anticipando la respuesta a la situación violenta que pueda desarrollarse al advertir, de las emociones monitorizadas, un contexto de riesgo próximo.

Además, en el marco del proyecto se pretende analizar la idoneidad de los protocolos de atención a las víctimas, especialmente en cuanto a los mecanismos preventivos para, a partir de los testimonios de estas, proponer nuevos criterios, atendiendo a las carencias expuestas por las víctimas y, en todo caso, auspiciando la inclusión de la tecnología (SAN SEGUNDO y LÓPEZ ONGIL, 2022, p. 22).

Por otro lado, destaca el avance en el proyecto IMPROVE, a través de la creación de un robot conversacional que posibilite una comunicación con las víctimas y que, mediante el recurso a sistemas de inteligencia artificial, permita ofrecer respuestas sobre sus derechos, valorar la situación de riesgo ante la que se encuentre, informar de recursos disponibles, entre otros.

Al respecto, se ha señalado la intención de que en el futuro las propias usuarias puedan descargarlo en sus dispositivos y les atienda cual agente, dando la información pertinente y respondiendo a sus cuestiones, aspecto este que, por sí solo, no parece resultar positivo si el objetivo del proyecto supone la sustitución del profesional y no el *mero* auxilio.

Por tanto, se aprecia la ventaja de contar con una herramienta preparada para auxiliar a la víctima en torno a derechos, prestaciones, teléfonos de contacto e incluso la llamada a terceros directamente a través del robot conversacional que, incluso, pueda identificar una posible situación de violencia y *recomiende* a la usuaria acudir a los profesionales pero, de pretender la descarga de una aplicación a modo de chatbot que pueda desplazar al profesional en este ámbito, lo cierto es que no resultará adecuado por cuanto el acompañamiento de la víctima es una constante petición por parte de las asociaciones de víctimas, conscientes de las necesidades de asistencia –de muy diversos tipos– que requieren en términos generales ante las situaciones de violencia padecida.

Señalados estos ejemplos, conviene atender los mismos con precaución. Si bien es loable cualquier tentativa de mejora en cuanto a la prevención se refiere, lo cierto es que el uso de estas herramientas expuestas puede conllevar diversas problemáticas asociadas que podrían resultar contraproducentes precisamente con los objetivos perseguidos, como las que se presentan a continuación.

Por ello, hay que reiterar que, en ningún caso, el desarrollo de estas tecnologías puede suponer el desplazamiento de los profesionales, más en un contexto donde se ha enfatizado insistentemente la necesidad de personal formado y acompañamiento de las víctimas. En todo caso, las herramientas deberán ser un auxilio y no una sustitución del factor humano. Solo los profesionales sabrán realmente tomar decisiones acertadas atendiendo a las circunstancias del caso, de la víctima, del agresor y de los condicionante--tan diversos-- que puedan darse entre los casos que se les presenten, con el recurso a estos sistemas que puedan servirles de ayuda, pero nunca sustituyendo la formación y la experiencia como solo una persona puede desarrollar.

En caso contrario, los sistemas apuntados anteriormente actuarán únicamente en los supuestos donde las víctimas encajen en el patrón advertido --sea en la monitorización de emociones, en el patrón de voz o ante un robot conversacional, entre otros-- pero dejarán de lado a aquellas que puedan no encontrarse bajo lo que estos sistemas hayan entendido por *víctima*.

Es otras palabras, los sistemas planteados tienen los datos ofrecidos por parte de las víctimas que, hasta el momento, se han investigado. Sin embargo, el fenómeno de la violencia contra las mujeres tiene sus particularidades, las cuales imposibilita establecer un perfil definido de víctima, siendo estas un conjunto heterogéneo de personas de las cuales, el único elemento común, es que son mujeres, pero sin posibilidad de hacer extensible las características de una, a las de todas; ni la respuesta emocional de una mujer ante una situación, a la respuesta que pueda tener otra mujer ante la misma situación (ni siquiera la misma, pues puede encontrarse ante otros condicionantes que varíen su respuesta).

Con ello, resulta complejo asegurar que una herramienta pueda ser utilizada en todas las manifestaciones de esta violencia, frente a cualquier víctima y en cualquier situación. Por tanto, si estos recursos se utilizan en lugar de los profesionales o si estos se atienen a sus resultados de forma obcecada, las consecuencias que pueden derivarse es la existencia de ciertas víctimas que no se inserten en los perfiles de estas herramientas y que, por tanto, en lugar de contar con herramientas que apoyen la prevención, nos encontremos con sistemas que desvirtúen el objetivo, excluyendo institucionalmente a mujeres víctimas por no acomodarse en el perfil de víctima que ha definido el sistema.

Por último, otra herramienta que destaca en este contexto es RAIST (Red flag artificial intelligence for sexual trafficking) creada por DIACONIA (siendo esta una plataforma de acción social que cuenta con diversos proyectos, entre otros uno encaminado en la lucha contra la trata de personas). Esta herramienta se basa en un modelo algorítmico que posibilita la detección temprana de captación de víctimas de trata, especialmente con fines de explotación sexual, a través de la herramienta Telegram.

En este sentido, dada la evolución tecnológica, se ha advertido cómo la criminalidad también adapta sus técnicas a las posibilidades que las TIC ofrecen. Así, se observa que en la actualidad las plataformas de comunicación, así como las redes sociales, se establecen como escenario perfecto para la captación de víctimas, especialmente en edades tempranas.

Por ello, esta herramienta destaca al haber advertido los peligros de las redes sociales y las plataformas de comunicación pero, a su vez, haber establecido herramientas preventivas a través de estas, si bien inicialmente solo a través de Telegram, con el objetivo de ampliarse a otras aplicaciones, una vez el proyecto piloto se asiente. Por medio de este sistema, analizando los patrones conversacionales que establecen habitualmente los captadores, se permite alertar al usuario de la posible situación de captación y sugerirle el contacto con un profesional en caso de que así lo estime necesario.

Con ello, esta herramienta es ejemplar por diversas cuestiones. Entre otras, porque aborda la prevención de la violencia contra las mujeres estableciendo medidas ajenas al ámbito de la pareja –donde normalmente se circunscriben todos los recursos– por otro, porque no sustituye a los profesionales, sino que permite que la herramienta sea un trampolín hacia ellos; por último, porque el mismo escenario que posibilita mayor victimización, es utilizado para advertir patrones de riesgo y para prevenir la captación.

Una vez que se ha hecho una aproximación general a diversas herramientas planteadas en esta materia, así como aquellos aspectos problemáticos de estas, conviene plantear aquí una propuesta en este campo, como es del aprovechamiento de las numerosas ventajas de las herramientas de IA para que estas sean desarrolladas, en esta materia, en un contexto preventivo en etapas educativas dado que los datos muestran que la violencia entre la población joven se sigue produciendo y se transforma, a medida que la tecnología avanza, desarrollándose una ciberviolencia que debe ser atendida especialmente, dadas sus particularidades.

Quizás el uso de esa tecnología para elementos preventivos en la adolescencia, así como para la identificación temprana de violencia, es un campo por explorar que debe ser considerado, valorando las capacidades que impregna la IA para una aplicación en estadios iniciales de esta violencia, abordando cuestiones como las relaciones de pareja, el amor romántico, la educación sexo-afectiva, la masculinidad hegemónica, la transversalidad de la violencia, la ciberviolencia, entre otras.

En definitiva, abordando numerosos aspectos vinculados con comportamientos violentos en esta materia que puedan ser prevenidos en etapas educativas, donde poner el foco no solo en que las víctimas puedan advertir sus manifestaciones en sus primeras fases –las más *invisibles*– sino especialmente en el tratamiento preventivo de posibles agresores futuros –cuando no ya presentes– todo ello recurriendo a la IA lo que permitiría, no solo un abordaje más eficiente –siempre acompañado de profesionales– sino una aproximación a la población joven más eficaz, por cuanto su vinculación con la tecnología es tal que, guste o no, cualquier atisbo de prevención y tratamiento también debe contar con ella.

Con todo y en conexión con lo señalado anteriormente, si bien es loable cualquier avance en el ámbito de la prevención y el tratamiento, lo cierto es que también deben apreciarse aquellas cuestiones que puedan ser contraproducentes, como los riesgos observados en el uso de herramientas de inteligencia artificial, como son el posible reforzamiento de desigualdades tradicionalmente asentadas y el (re)asentamiento de una violencia simbólica, especialmente protagonista en el ámbito de la violencia contra las mujeres, tal y como se expondrá a continuación.

3. Reforzamiento de desigualdades y violencia simbólica: ¿peaje frente al desarrollo tecnológico?

Como se ha podido apreciar en líneas precedentes, el uso de herramientas de IA no constituye un futurible, sino que se establece como realidad presente en multitud de ámbitos, entre los que se encuentra el fenómeno de la violencia contra las mujeres.

Si bien se valoran las muy diversas ventajas de su aplicación, especialmente la realización de tareas en mucho menor tiempo de lo que lo haría una persona o la posibilidad de establecer predicciones a través de operaciones matemáticas en escaso tiempo, con menor error y con absoluta destreza, lo cierto es que los aspectos negativos advertidos también han protagonizado los análisis en este campo.

Entre otros, cabe destacar cómo la propia UNESCO en su Recomendación sobre la ética de la inteligencia artificial, alude, en su considerando, a las preocupaciones éticas que suscitan con su desarrollo⁶, señalando

en relación con los sesgos que pueden incorporar y exacerbar, lo que puede llegar a provocar discriminación, desigualdad, brechas digitales y exclusión y suponer una amenaza para la diversidad cultural, social y biológica, así como generar divisiones sociales o económicas; la necesidad de transparencia e inteligibilidad del funcionamiento de los algoritmos y los datos con los que han sido entrenados; y su posible impacto en, entre otros, la dignidad humana, los derechos humanos y las libertades fundamentales, la igualdad de género, la democracia, los procesos sociales, económicos, políticos y culturales, las prácticas científicas y de ingeniería, el bienestar animal y el medio ambiente y los ecosistemas

En este sentido, diversas investigaciones han demostrado la perpetuación de sesgos y estereotipos tradicionales que fortalecen situaciones discriminatorias a través del uso de herramientas de IA, cuestionando el respeto a derechos fundamentales, especialmente vinculados con la igualdad de las personas, donde es precisamente en el campo de la violencia contra la mujer donde esta cuestión debe ser tenida especialmente en cuenta.

Así, cuando nos referimos a este tipo de violencia, tendencialmente sobresalen aquellas manifestaciones más graves de violencia física, dejando en un segundo plano el grueso del fenómeno, constituido por otro tipo de hechos que, en todo caso, son sustentados por patrones culturales propios de una sociedad patriarcal.

En este sentido, los estereotipos y sesgos advertidos en los estudios han señalado que aquellos relativos al sexo y al género se encuentran impregnando los más sofisticados recursos de inteligencia artificial –aunque también lo hacen otros, como los vinculados con la etnia o la clase– (como señala FERRANTE, 2021).

De esta forma, los algoritmos que se han ido expandiendo en la sociedad, lejos de evitar errores humanos y promover decisiones más adecuadas, se traducen en resultados no neutrales (en la línea de lo que exponen SMITH, G. Y RUSTAGI, 2021), produciendo lo que se ha denominado como “sesgo algorítmico”.

⁶ Disponible en: https://unesdoc.unesco.org/ark:/48223/pf0000380455_spa

De esta forma, ARÁNGUEZ (2022, p. 73 y ss.) señala cómo los sesgos sexistas presentes en los algoritmos se traduce en una reproducción de la discriminación presente en los datos que incorporan las herramientas de IA, así como alimentan el sexismo de los usuarios de estos sistemas.

Entre otros ejemplos prácticos, se ha analizado cómo en las herramientas de prevención delincriminal, así como las relativas a la toma de decisiones sobre suspensiones de condena o beneficios penitenciarios, estas han reforzado representaciones discriminatorias, siendo más exigentes con un tipo de sujetos (especialmente con aquellos racializados y/o de clases sociales más empobrecidas).

En este sentido, defiende acertadamente Miro Llinares (2018) que en un contexto como el de la criminalidad, nos encontramos ante “demasiadas aristas como para una predicción objetiva y libre de sesgos” (p.96), lo cual cuestiona la posibilidad de aplicación de este tipo de técnicas en estos ámbitos, donde las garantías deben ser especialmente cuidadas y donde las decisiones tomadas condicionan significativamente la vida de las personas, por lo que no resulta aceptable mantener instrumentos con análisis despersonalizados.

Otro ejemplo, siendo un recurso muy extendido, es el del reconocimiento facial, donde se ha advertido mayor imprecisión con determinados grupos de personas (funcionando con mayor rigor cuando los sujetos eran hombres caucásicos) y siendo menos precisos cuando los sujetos son racializados o pertenecen a categorías de género no dicotómicas (como defiende Ortiz De Zarate, 2023).

Por otro lado, se han observado resultados discriminatorios en ámbitos como el de la traducción de textos, donde idiomas neutrales al género se han traducido, de forma sexista, al convertirlos en idiomas donde sí se producen modificaciones conforme al género. Aquí podría ejemplificarse con la traducción de “doctor”, como médico y “nurse” como enfermera (como señala en su estudio PARSHEERA, 2018) respondiendo a una tradicional distinción de roles en el ámbito laboral que no hacen sino perpetuar la segregación horizontal.

En este sentido, también se ha observado cómo la inclusión de algoritmos en el ámbito de los recursos laborales ha supuesto peligrosas exclusiones para las mujeres. Así, la empresa Amazon había apostado por el recurso a la IA para mejorar el sistema de incorporación de trabajadores/as a su empresa. Sin embargo, el algoritmo tendía a dejar de lado los CV de las mujeres, por el mero hecho de serlo, pese a disponer de las cualidades necesarias para el puesto (como desarrolla Dastin, 2018).

En esencia, el algoritmo se creaba partiendo de una cantidad de datos previos que se correspondían, esencialmente, a trabajadores hombres. Al crear un patrón de búsqueda para una persona que encajara en el puesto, conforme a los trabajadores previos, el resultado producido era la penalización –y, por tanto, exclusión– de los CV de mujeres. Con ello, si bien el sistema no estaba creado con ningún objetivo discriminatorio, ni había recibido órdenes en este sentido, los efectos producidos sí lo eran.

Asimismo, en la línea de apreciar cómo los avances tecnológicos tienden a perpetúan roles tradicionales, es posible destacar, como ejemplo, la creación de asistentes virtuales con voces de mujeres –conservando el rol de mujer como “servicial”– e incluso la asigna-

ción de nombres “femeninos” a robots que están creados para tareas típicamente femeninas, como las de la limpieza; mientras que se asignan nombres “masculinos” a tareas tradicionalmente vinculadas con los hombres (como indica Delicado-Moratalla, 2021).

En esta línea, se señala que la proliferación de asistentes virtuales con voces femeninas puede transmitir –y, por tanto, perpetuar– la idea de que las mujeres son serviciales y dóciles; mientras que se apuesta por el mantenimiento de voces masculinas en determinados servicios, aludiendo a que el consumidor lo prefiere cuando se requieren respuestas autoritarias o explicaciones “complejas”, tal y como señala la UNESCO (2019) en su informe *I'd blush if I could*. Así, estudios como el de ACOSTA-CALDERÓN Y BENITEZ-EYZAGUIRRE (2022) confirman que existe un “diseño sexista” de este tipo de dispositivos donde, además, el usuario no es consciente de este sesgo (p.35)

En este ámbito, merece la pena destacar la creación de los denominados *robots sexuales*, los cuales se definen como una “muñeca hinchable hiperrealista con inteligencia artificial” (DELICADO-MORATALLA, 2021, p. 221), que promueven riesgos que deben ser atendidos. En concreto, con la creación de estos robots se están desarrollando nuevos espacios que propician la denominada *cultura de la violación* favoreciendo que, a través de las robots, se mantenga la cosificación de la mujer y se representen situaciones de dominación, sometimiento y violencia sexual.

Con ello, nos encontramos ante un momento en el que conviven dos realidades contrapuestas. Por un lado, aquella que advierte el problema de la violencia sexual y reconoce la necesidad de atenderlo potenciando una adecuada salud sexual-afectiva desde las etapas escolares, tal y como defiende la Ley Orgánica 10/2022, de 6 de septiembre, de garantía integral de la libertad sexual, así como la Ley Orgánica 1/2023, de 28 de febrero, por la que se modifica la Ley Orgánica 2/2010, de 3 de marzo, de salud sexual y reproductiva y de la interrupción voluntaria del embarazo; entre otras. Por otro lado, la del pernicioso avance tecnológico que dinamiza nuevas formas de antiguas violencias.

En este contexto, nos encontramos en un momento de sensibilización frente a la necesidad de una adecuada educación en sexualidad desde etapas tempranas, dado que se ha advertido una ausencia de educación por parte de escuelas y familias que, por el contrario, se ve sustituida por el recurso a la pornografía –iniciándose en edades muy bajas, siendo los 8 años la edad de primera visualización de contenido pornográfico, encontrándose con un consumo habitual en los 12–.

Esto supone que los jóvenes normalicen relaciones sexuales irreales y, en muchos casos, constitutivas de violencia sexual, más que de una relación sexual consentida donde, en cualquier caso, generalmente tiende a posicionarse a la mujer en situaciones de sometimiento y vejaciones.

Con ello, el de las relaciones sexuales suele ser un contexto donde la expresión del patriarcado se presenta con mayor intensidad y donde se está apreciando cómo una nefasta educación, acompañada del resto de conductas que promueven la cultura de la violación, se traduce en el incremento de agresiones sexuales producidas por población joven –cuando no menor de edad–.

Por ello, se cuestiona especialmente cómo el avance tecnológico y el recurso de la IA desarrolle nuevos mecanismos de reproducción de la violencia sexual que dé cabida a nuevas prácticas y nuevos espacios de impunidad.

En definitiva y, sin posibilidad de mayor profundización por las limitaciones del texto, se debe admitir que los algoritmos, lejos de ser neutrales, están repletos de valores en todas sus fases (sea en la identificación del problema, en la selección de datos e incluso en la elección del algoritmo) lo cual puede reafirmar un contexto discriminatorio, especialmente vinculado con los prejuicios de género derivándose esencialmente de la sobrerrepresentación masculina en el desarrollo de la inteligencia artificial, además de la ausencia de la perspectiva de género en este campo, como señalan Cernadas y Calvo (2022).

Esto, en el ámbito de la violencia contra las mujeres es especialmente preocupante, por cuanto esta se nutre de una cultura patriarcal discriminatoria que, pese a que parezca paradójico, se afianza con los avances tecnológicos, lo cual posibilita la perpetuación del sustrato de esta violencia y que cualquier tratamiento devenga en fracaso.

No obstante, no debe entenderse como *peaje* ante el desarrollo tecnológico, ni negar el avance en este campo. Por el contrario, debe partirse de la aceptación de encontrarnos ante una realidad “con alto riesgo discriminatorio por naturaleza” (Instituto de las Mujeres, 2023, p. 12), para conseguir erradicarla.

Entre otras soluciones, se plantean el establecimiento de equipos multidisciplinares en todas las fases de creación de las herramientas de IA, así como sistemas de monitorización periódica, que permitan advertir en el menor tiempo posible cualquier problema derivado del algoritmo.

De esta forma, si bien la erradicación total de cuestiones discriminatorias pueda ser un pensamiento utópico, al menos las posibilidades serán reducidas conforme los objetivos de los algoritmos tengan voluntades más igualitarias, el personal que los desarrolle sea heterogéneo y formado en perspectiva de género –y en materia de igualdad en su sentido amplio– y se cuenten con sistema de monitorización permanente que identifique cualquier desvío inadecuado de la herramienta, para ser convenientemente corregido antes de que las consecuencias sean más perjudiciales.

4. Conclusiones y recomendaciones en pro de una inteligencia artificial comprometida con la igualdad

Tal y como se apuntaba inicialmente, pese a la ausencia de una conceptualización unívoca sobre la inteligencia artificial, lo cierto es que ha alcanzado tal protagonismo en la actualidad que se han desarrollado numerosas investigaciones en torno a sus posibilidades. Se han analizado cuestiones éticas vinculadas con su uso, así como problemas prácticos en su aplicación, esencialmente vinculados con las garantías procesales o el respeto a los derechos fundamentales, pero también en conexión con el recurso a la inteligencia artificial para una *modernización criminal* ante la que nos encontramos, por cuanto los avances en la materia también son y serán utilizados para la comisión delictiva, permitien-

do mayor alcance, variaciones delictivas, menor posibilidades de identificación criminal, entre otras cuestiones que deben ser atendidas por el derecho penal.

En este contexto, se ha destacado el empleo de la inteligencia artificial en un ámbito como es el de la violencia contra las mujeres, donde la aplicación de algoritmos tiene ya cierta vigencia, esencialmente a través del uso del sistema VIOGEN. Este sistema, con un objetivo loable y diversos seguimientos para su mejora, sigue planteando cuestionamientos relevantes. Entre otros, la automatización en la respuesta de los profesionales que se encuentran detrás, aceptando *ciegamente* el resultado obtenido por parte del algoritmo, cuya valoración de riesgo escasas veces se ve modificado.

Como se ha apuntado, al sistema VIOGEN le siguen en la actualidad numerosos impulsos en pro del uso de inteligencia artificial en este ámbito. Si bien no con la misma relevancia, sí se aprecian interesantes avances en la materia, con el desarrollo de robots conversacionales, sistemas de monitorización de emociones o patrones de voz que permitan advertir contextos de violencia, enmascarados en patologías “comunes”.

Si bien esta expansión debe ser aplaudida, pues debe valorarse cualquier avance en materia de prevención de este tipo de violencia, lo cierto es que no deben pasarse por alto los riesgos derivados del uso de este tipo de herramientas, especialmente ante víctimas que se encuentran en una situación especialmente vulnerable. Entre otras, tal y como se ha apuntado, la heterogeneidad de las víctimas imposibilita herramientas únicas, con capacidad de abordar toda manifestación de violencia. Estandarizar las respuestas frente a conductas tan variadas, en mujeres muy diversas, podría hacer que cierto tipo de mujeres no *encaje* con los patrones institucionalizados y que, por ello, se vean expulsadas del sistema, obteniendo, por tanto, muy perniciosas consecuencias, contrarias completamente a los objetivos previstos.

Asimismo, se ha observado cómo el empleo de inteligencia artificial impregna muy diversos contextos, como aquellos vinculados con la economía, el ámbito laboral, la justicia, entre otros. Esto ha desarrollado numerosos debates, por cuanto se ha advertido el peligro de la extensión de una moderna discriminación, escondida en sistemas supuestamente neutrales, que, en la práctica, se basan en sesgos marcadamente tradicionales.

Esta cuestión, vinculada con el tema que ocupa, es especialmente perniciosa, dada la base cultural patriarcal que sustenta la violencia contra las mujeres, que podría ser reforzada gracias a sistemas de inteligencia artificial aparentemente neutrales, pero que enmascaren patrones culturales marcadamente sexistas y discriminatorios con las mujeres. Por tanto, encontrándonos ante avances tecnológicos que escondan un claro retroceso en materia de igualdad, fortaleciendo contextos discriminatorios y propiciando una peligrosa violencia simbólica.

Con todo, no debe demonizarse el recurso a la inteligencia artificial ni, en concreto, el uso de algoritmos, por cuanto pueden suponer numerosas ventajas que ayuden en la toma de decisiones, en la gestión de datos y en el ahorro de tiempo y esfuerzo. Todo lo contrario, debe advertirse sus beneficios, pero no de forma obnubilada, sino atendiendo a los riesgos que presenta y valorando la forma de limitarlos.

Asimismo, se debe luchar contra un determinismo algorítmico, que obvie la importancia humana. Es decir, si bien las ventajas de la inteligencia artificial son numerosas, en ningún caso deben servir para sustituir las decisiones humanas, sino como *mero* auxilio en la toma de decisiones, especialmente cuando nos encontramos ante un contexto como el de la violencia contra las mujeres.

Por último, en este campo se debe plantear, como línea de futuro, que el uso de la inteligencia artificial en materia de violencia contra la mujer se desarrolle en los escenarios más iniciales, es decir, como mecanismo de identificación de esta violencia en etapas escolares y promoción de la formación en estas edades, donde la violencia se sigue produciendo y reinventando, especialmente con las TIC –encontrándonos ante la ciberviolencia de género– y donde los recursos siguen siendo escasos.

Quizás sea en este escenario donde las investigaciones en materia de IA y de sus posibilidades deberían continuar avanzando.

Bibliografía

- Acosta-Calderón y Benitez-Eyzaguirre (2022). Siri, desde la perspectiva feminista. En Aránguez Sánchez, T. y Olariu, O. (Cord.) *Algoritmos, teletrabajo y otros grandes temas del feminismo digital* (pp. 21-37) Dykinson.
- Aránguez Sánchez, T. (2022) sesgos sexistas de los algoritmos e inteligencia artificial. En Aránguez Sánchez, T. y Olariu, O. (Cord.) *Algoritmos, teletrabajo y otros grandes temas del feminismo digital* (pp. 71-88) Dykinson.
- Borges Blázquez, R. (2020), “El sesgo de la máquina en la toma de decisiones en el proceso penal”, *IUS ET SCIENTIA*, 2.
- Cernadas, E. y Calvo, E. (2022), “Perspectiva de género en inteligencia artificial, una necesidad”, *cuestiones de género: de la igualdad y la diferencia*, nº17, pp. 111-127.
- Dastin, J., “Amazon scraps secret AI recruiting tool that showed bias against women”, *reuters*, 11 de octubre de 2018.
- Delgado Calvo; M. (1996), *La inteligencia artificial. Realidad de un mito moderno*, Granada.
- Delicado-Moratalla, L. (2021), “La robot sexual y la pornografía: la ilusión del poder masculino y la fantasía de cosificar a las mujeres”, *Atlánticas. Revista Internacional de Estudios Feministas*, nº 6.
- Del Rosal Blasco, B. (2023). ¿El modelo de la responsabilidad penal de las personas jurídicas para los daños punibles derivados del uso de la inteligencia artificial?, *revista de responsabilidad penal de personas jurídicas y compliance*, vol. 2, pp. 1-49.
- Fernández, A. (2019) Inteligencia artificial en los servicios financieros, *Boletín económico del Banco de España*, 2.
- Ferrante, E., (2021) “Inteligencia artificial y sesgos algorítmicos ¿Por qué deberían importarnos?”, *revista nueva sociedad*, nº 294.
- Hernández Giménez, M. (2019), “Inteligencia artificial y derecho penal”, *Actualidad Jurídica Iberoamericana*, nº 10 bis, pp. 792-843.

- Instituto de las Mujeres (2023). Informe preliminar con perspectiva interseccional sobre sesgos de género en la inteligencia artificial. Disponible en: https://www.inmujeres.gob.es/areasTematicas/SocInfo/Estudios/docs/Informe_Sesgos_Genero_IA.pdf
- Ley Orgánica 10/2022, de 6 de septiembre, de garantía integral de la libertad sexual.
- Ley Orgánica 1/2023, de 28 de febrero, por la que se modifica la Ley Orgánica 2/2010, de 3 de marzo, de salud sexual y reproductiva y de la interrupción
- Miró Llinares, F. (2018), "Inteligencia artificial y justicia penal: más allá de los resultados lesivos causados por robots", *revista de derecho penal y criminología*, 20, pp.87-130
- Miró Llinares, F. (2020), "El sistema penal ante la inteligencia artificial: actitudes, usos, retos", en Dupuy de Repetto y Corvalán (dir.): *Cibercrimen III: inteligencia artificial, automatización, algoritmos y predicciones en el derecho penal y procesal penal*, Buenos Aires, pp. 81-136.
- Morillas Fernández, D. L. (2023). Implicaciones de la inteligencia artificial en el ámbito del Derecho Penal, *Derecho Penal, Inteligencia Artificial y Neurociencias*. Roma Tre-Press, pp. 59-91.
- Muñoz Rodríguez, A.b. (2020), "El impacto de la inteligencia artificial en el proceso penal", *Anuario de la Facultad de Derecho. Universidad de Extremadura*, 36, p. 699.
- Ortiz De Zarate, L. (2023), "Sesgos de género en la inteligencia artificial", *revista de Occidente*, nº502, pp. 5-20.
- Parsheera, S. (2018), "A gendered perspective on artificial intelligence. proceedings of itu kaleidoscope 2018 — machine learning for a 5G Future", *IEEE*.
- Real Decreto 729/2023, de 22 de agosto, por el que se aprueba el Estatuto de la Agencia Española de Supervisión de Inteligencia Artificial.
- Rodríguez-Gordo, C. y Sarrión-Andaluz, J. (2021), Polarización social y control político: algunas consecuencias de la inteligencia artificial y las redes sociales para la razón práctica, *Cuadernos Salmantinos de Filosofía*, vol.48, pp. 157-185.
- Sánchez Rosado, J. C. y Díez Parra, M. (2022). Impacto de la inteligencia artificial en la transformación de la sanidad: beneficios y retos, *Economía industrial*, (423), pp. 129-144.
- San Segundo, R. y López Ongil, C. (2022). EMPATIA-CM: Comprehensive Protection of Gender-based Violence Victims through Multimodal Affective Computing, *Sociología y tecnología*, 12, pp. 1-33.
- Simón Castellano, P. (2021). Inteligencia artificial y Administración de Justicia: ¿Quo vadis, justitia?, *IDP revista de Internet, Derecho y Política*, 33.
- SMITH, G. y RUSTAGI, I., "When Good Algorithms Go Sexist: Why and How to Advance AI Gender Equity", *Stanford social innovation review*, 31 de marzo de 2021.
- UNESCO (2019). I'd blush if i could. Closing gender divides in digital skills through education. Disponible en: <https://en.unesco.org/Id-blush-if-I-could>
- Valls Prieto, J. (2021). *Inteligencia artificial, Derechos humanos y bienes jurídicos*. Aranzadi.
- Zurita (2014). *Violencia contra la mujer. Marco histórico evolutivo y predicción del nivel de riesgo* [Tesis doctoral]. Universidad Autónoma de Madrid.