



I+G 2022

## VIII Congreso Universitario Internacional Investigación y Género 2022

23 y 24 de junio de 2022  
Universidad de Sevilla

# INVESTIGACIÓN Y GÉNERO Proyectos y Resultados en Estudios de las Mujeres

María Elena García-Mora y Ana María De la Torre-Sierra (Eds.)



Seminario Interdisciplinar  
de Estudios de las Mujeres

Universidad de Sevilla  
2022

VIII Congreso de Investigación y Género. Reflexiones sobre investigación para avanzar en igualdad.

Universidad de Sevilla, 2022.

Reservados todos los derechos. El contenido de esta obra está protegido por la ley, que establece penas de prisión y/o multas, además de las correspondientes indemnizaciones por daños y perjuicios, para quienes reprodujeren, plagiaran, distribuyeren o comunicaren públicamente, en todo o en parte, una obra literaria, artística o científica, o su transformación, interpretación o ejecución artística fijada en cualquier tipo de soporte o comunicada a través de cualquier otro medio, sin la preceptiva autorización.

I.S.B.N: 978-84-09-41805-3

# JUSTICIA E INTELIGENCIA ARTIFICIAL EN CLAVE DE GÉNERO<sup>1</sup>

Arruti Benito, Sara<sup>2</sup>

## CUESTIONES CONCEPTUALES PREVIAS: INTELIGENCIA ARTIFICIAL, ALGORITMOS Y MACHINE LEARNING

### *Contextualización y significado de la Inteligencia Artificial*

El imparable avance tecnológico ha generado una verdadera metamorfosis en nuestra forma de vivir, de trabajar y de relacionarnos los unos con los otros. En este contexto, hemos asistido a la denominada 4.0 Revolución Industrial caracterizada por la aparición y la fusión de nuevas tecnologías –el *blockchain*, el *big data*, la robótica...– que están eliminando las líneas entre lo físico, lo digital y lo biológico<sup>3</sup>. En esta nueva sociedad completamente digitalizada, se ha ido gestando lo que conocemos como Inteligencia Artificial<sup>4</sup> (en adelante, IA).

En cuanto a terminología se refiere, definir la IA no es tarea fácil, debido al carácter complejo, evolutivo y variable de las tecnologías que la configuran. No obstante, en el ámbito del Consejo de Europa, el Grupo de Expertos de Alto Nivel en IA<sup>5</sup> ha consensuado una definición de los sistemas de IA como “...sistemas de *software* (y posiblemente también *hardware*) diseñados por humanos que, dado un objetivo complejo, actúan en la dimensión física o digital, percibiendo su entorno a través de la adquisición de datos, interpretando datos estructurados o no estructurados recopilados, razonando sobre el conocimiento o procesando la información, derivada de estos datos, y decidiendo cuál es la mejor acción a tomar para lograr un objetivo dado”<sup>6</sup>. En este sentido, aunque no exista aquiescencia sobre la definición de este fenómeno, se divisa un denominador común que convierte la IA en una combinación compleja de tecnologías que agrupan datos, algoritmos y capacidad informática con el objeto de intentar emular en funcionamiento de la inteligencia humana<sup>7</sup>. Es

<sup>1</sup> La presente comunicación está vinculada a la Tesis Doctoral “Análisis de la perspectiva de género en la estructura y en la actividad del Poder Judicial”, bajo la dirección y supervisión de Ixusko Ordeñana Gezuraga, Profesor Titular de Derecho Procesal de la Universidad del País Vasco (UPV/EHU).

<sup>2</sup> Doble graduada en Derecho y Administración de Empresas y Doctoranda del programa “Derechos Humanos, Poderes Públicos, Unión Europea: Derecho Público y Privado” de la Universidad del País Vasco, sarruti002@ikasle.ehu.eus

<sup>3</sup> A pesar de que aparentemente nos encontramos ante una disciplina nueva y novedosa, se sostiene que el concepto de inteligencia artificial data de 1955, cuando John McCarthy y su grupo, Marvin Minsky, Nathaniel Rochester y Claude Shannon idearon un proyecto que se basaba en la posibilidad de crear máquinas que pudieran realizar tareas a través de una programación previa que imitara las redes neuronales de forma similar al pensamiento humano. Véase, Solar Cayón, José Ignacio (2019): *La inteligencia artificial jurídica. El impacto de la innovación tecnológica en la práctica del Derecho y el mercado de servicios jurídicos*, Aranzadi, Pamplona.

<sup>4</sup> Para mayor profundización sobre la complejidad y la trascendencia de la 4.0 Revolución Industrial, véase Schwab, Klaus (2017): *The Fourth Industrial Revolution*, Crown Business Publishing, Nueva York.

<sup>5</sup> La Comisión Europea creó en 2018 un grupo de Expertos de Alto Nivel en IA que tiene como objetivo general apoyar la aplicación ética de los sistemas de IA. Este grupo de expertos y expertas entienden que los Derechos Fundamentales son la columna vertebral de la programación de estos sistemas.

<sup>6</sup> Comisión Europea (2019). *A definition of AI: Main capabilities and scientific disciplines*, Bruselas, p.6. Documento electrónico, <https://www.aepd.es/sites/default/files/2019-12/ai-definition.pdf>, consultado: 18/04/2022.

<sup>7</sup> Comisión Europea (2019). *A definition of AI: Main capabilities and scientific disciplines*, Bruselas, p.6. Documento electrónico, <https://www.aepd.es/sites/default/files/2019-12/ai-definition.pdf>, consultado: 18/04/2022.

decir, un sistema de IA –en términos generales– es un algoritmo que hace recomendaciones, predicciones o adopta decisiones para un conjunto de objetivos previamente programados<sup>8</sup>. Por tanto, los sistemas de IA dependen, por un lado; de un algoritmo configurado de antemano y, por otro lado; de un extenso conjunto de datos de datos que nutre dicho sistema.

### ***Funcionamiento de los algoritmos y el Machine Learning***

En esta línea, resulta determinante conocer someramente el funcionamiento y desarrollode un algoritmo para poder comprender su alcance discriminatorio. Los algoritmos son series de ecuaciones matemáticas que se van entrelazando para proporcionar un resultado<sup>9</sup>. Su funcionamiento se basa en el empleo aleatorio de un gran conjunto de datos (*Smart data*) que, bajo un modelo matemático, establecen patrones de correlación entre ellos. Estos patrones decorrelación son lo que permiten al algoritmo realizar apreciaciones, predicciones y adoptar decisiones sobre el fin para el que se ha programado. Por tanto, la finalidad de un algoritmo no es adoptar una decisión justa, equitativa, objetiva y verdadera, sino establecer correlaciones, buscar patrones lógicos, entre los datos ofrecidos.

En este contexto, cobra especial relevancia el *Machine Learning* o los sistemas de aprendizaje automático, por el particular impacto discriminatorio que puede derivar de su funcionamiento y de su aplicación en la práctica judicial. Estos sistemas son algoritmos muy complejos y sofisticados que tienen la capacidad de aprender de sí mismos. Pueden ser increíblemente eficaces a la hora de predecir comportamientos y establecer patrones, no obstante; dicha eficacia confronta con un problema de difícil solución: su funcionamiento evoluciona hasta desarrollar una inteligencia autónoma y singular, incomprensible para la mente humana<sup>10</sup>. En esta línea, se deben mencionar lo que se conoce como cajas negras o *black boxes*<sup>11</sup>. Este tipo de sistemas, a diferencia de las cajas blancas, no permiten destripar el algoritmo y comprender las decisiones que ha adoptado. Esta característica presenta un enorme desafío en cuanto a la garantía de los derechos fundamentales de las personas. Particularmente, desde el derecho y desde la perspectiva de género, como remarcaremos, se requiere una mayor atención a estos sistemas de cajas negras, dado que pueden convertirse en herramientas con un potencial discriminatorio enormemente alto contra las mujeres, con la agravante de que dicho trato no pueda explicarse.

---

<sup>8</sup> Conviene puntualizar que, con el objeto de centrarnos en el debate jurídico, dejaremos a un lado las precisionestécnicas y utilizaremos las palabras algoritmos y sistemas de IA para referirnos al mismo concepto.

<sup>9</sup> Comisario de Derechos Humanos del Consejo de Europa (2019): *Unboxing Artificial Intelligence: 10 steps to protect Human Rights*, p.24. Documento electrónico, <https://rm.coe.int/unboxing-artificial-intelligence-10-steps-to-protect-human-rights-reco/1680946e64>, consultado: 19/04/2022.

<sup>10</sup> Borges Blázquez, Raquel (2021): “El sesgo de la máquina en la toma de decisiones en el proceso penal”, *IUS ET SCIENTIA*, Vol.6, N°2, p.58.

<sup>11</sup> La noción de caja negra se refiere a la inexplicabilidad de los pasos a través de los cuales el sistema de IA ha llegado a una decisión. El algoritmo conforme va evolucionando y editándose a sí mismo, va tomando en cuenta diferentes variables y va asignando diferentes pesos a cada variable a medida que analiza su grado de importancia en los patrones que percibe. De esta manera, llega un momento en el que no somos capaces, incluidos quienes han diseñado el sistema, de reconocer los patrones que se han empleado para llegar al resultado. Véase, O’Donnel, Renata M. (2019): “Challenging racist predictive policing algorithms under the equal protection clause”, *New York University Law Review*, Vol. 9, N°3, p.551.

## INTELIGENCIA ARTIFICIAL EN CLAVE DE GÉNERO: BRECHA DIGITAL, SESGOS Y DISCRIMINACIÓN ALGORÍTMICA

Una vez esbozadas las líneas conceptuales sobre las que se basa la presente contribución, es necesario realizar una revisión crítica de los mismos desde la perspectiva de género. En este sentido, nos interesa hacer un análisis desde la experiencia de las mujeres, es decir, introduciendo la mirada y la experiencia del género femenino en este nuevo mundo algorítmico, para ver, detectar y corregir -en este orden- todas las formas en las que la IA puede contribuir a perpetuar y a reforzar la discriminación contra las mujeres<sup>12</sup>. Este análisis resulta aún más necesario, teniendo en cuenta nuestra tendencia a pensar que la tecnología, y en este caso los algoritmos, son instrumentos matemáticos caracterizados por la lógica, la objetividad y la neutralidad, produciendo, por ende, resultados presuntamente racionales e imparciales.

En primer lugar, es necesario recordar que el conocimiento y el quehacer humano registrado a lo largo de nuestra historia, se ha construido para y desde el hombre. Desde esta perspectiva androcéntrica<sup>13</sup>, tomando como paradigma de lo humano únicamente al varón, se han diseñado todos los vestigios del conocimiento humano (historia, arte ...), invisibilizando los deseos, necesidades y experiencias de la mitad de la humanidad, las mujeres. La Ciencia y la Tecnología, así como el Derecho, no se encuentran al margen de lo descrito.

En este contexto, es necesario remarcar que la IA se ha configurado en una sociedad que parte de una desigualdad estructural entre hombres y mujeres, donde predomina la subordinación de la mujer y las relaciones asimétricas de poder. Bajo este paradigma, la creación y diseño de los sistemas de IA no presentan un diseño neutral en términos de género. Dicha afirmación se advierte, por un lado, en la brecha digital de género existente en el ámbito de la IA, y, por otro, en la existencia de sesgos de género que dan lugar a una nueva modalidad de trato discriminatorio—la discriminación algorítmica por razón de género—.

### ***Brecha digital de género***

La IA es un campo donde predomina la participación y la mirada del género masculino. Muestra de ello son los últimos datos que arroja el “*Global Gender Gap Report 2021*”<sup>14</sup>, donde se establece que más del 75% de los profesionales de la IA son hombres. Esta brecha de género se engrosa si tenemos en cuenta la especialización por roles de género. Concretamente, en un estudio que ha realizado recientemente el Instituto de la Mujer, se ha puesto de manifiesto que las mujeres son relegadas a puestos de menor responsabilidad y menor retribución (análisis de datos, procesamiento de lenguaje natural y estructuras de datos, recuperación de información...). Muestra de ello es que apenas un 12% de las

---

<sup>12</sup> Puntualizamos que, aunque partamos de la experiencia de un sujeto específico, la perspectiva de género no pretende la centralidad del género femenino en el sentido de construir una mirada que se erija como única y aplicable a la totalidad de los seres humanos. El análisis de género, pone al descubierto como el hombre y lo masculino son el referente de la mayoría de las explicaciones de la realidad en detrimento de los derechos de las mujeres y de los valores asociados con lo femenino, y sugiere nuevas formas de construir los géneros que no estén basadas en la discriminación.

<sup>13</sup> Sobre la diferencia de la perspectiva androcéntrica y las género-sensitivas, véase, Facio, Alda, Fries, Lorena (1999): *Género y Derecho*, LOM ediciones, Santiago de Chile.

<sup>14</sup> World Economic Forum (2021): *Global Gender Gap Report 2021*, Suiza, p.60. Documento electrónico, <https://www.weforum.org/reports/global-gender-gap-report-2021>, consultado 20/04/2022.

personas que participan en el desarrollo de este tipo de innovaciones con capacidad de decisión de alto nivel sean mujeres o solo el 11% de las personas que programan código fuente (escriben *software*)<sup>15</sup>.

Estos datos ponen en evidencia que los sistemas de IA se diseñan con una infrarrepresentación del género femenino, desde una mirada unidimensional (masculina) que no abarca la diversidad de la sociedad, lo que contribuye, entre otras razones, a la aparición de sesgos de género que penetran en dichos sistemas.

### **La discriminación algorítmica de género: El problema de los sesgos.**

Para abordar la problemática de la discriminación algorítmica, se debe partir de la raíz de este fenómeno: la falsa neutralidad y objetividad de los algoritmos. Aunque suele presumirse que los sistemas de IA son neutros y deciden objetivamente, no debemos olvidar que la creación de estos sistemas parte de la acción humana<sup>16</sup>. La base de datos que sustenta todo sistema de IA, la crea una persona que se inserta en un determinado contexto social. Ello implica que estos algoritmos corran el riesgo de reflejar los sesgos –raciales, de clase, de género, de orientación sexual...– de aquellas personas que los diseñan<sup>17</sup>. En esta línea, requieren especial atención los sesgos de género, es decir, aquellas nociones preconcebidas o estereotipos que perjudican a una persona en función de su género, sea directa o indirectamente, dando lugar a resultados discriminatorios (en la mayoría de casos en contra de las mujeres). No obstante, la problemática no termina aquí, puesto que, como veremos *infra*, son diversas las fases y las razones por las que un sesgo perjudicial puede integrarse en los sistemas de IA y dar lugar a un trato discriminatorio, en este caso, en función del género.

#### *Fases donde se puede dar la discriminación algorítmica*

##### a) Momento inicial de delimitar el tema y definir el objetivo:

En un estadio inicial, antes de recoger y seleccionar los datos que nutrirán el algoritmo, es posible que los sesgos, en general, y los de género, en particular, interfieran en el momento inicial de determinar el tema y objeto para el que se programa el sistema de IA. Por ejemplo, en el ámbito jurídico, si queremos crear un algoritmo que tiene como propósito valorar la credibilidad de la declaración de una víctima de violencia de género en un proceso penal, tendremos que tener presente la abundante jurisprudencia del Tribunal Supremo en esta materia (en cuanto a la valoración del retraso en denunciar, la víctima como testigo cualificado, criterios de valoración de la declaración de aquella como única prueba de cargo...), para delimitar el tema y el objetivo adecuadamente.

##### b) Selección de los datos:

---

<sup>15</sup> Para mayor profundización véase, Instituto de la Mujer (2020): *Mujeres y digitalización. De las brechas a los algoritmos*, Catálogo de Publicaciones de la Administración General del Estado, Madrid, p. 59-60. Documento electrónico, [https://www.inmujeres.gob.es/disenovovedades/M\\_MUJERES\\_Y\\_DIGITALIZACION\\_DE\\_LAS\\_BRECHAS\\_A\\_LOS\\_ALGORITMOS\\_04.pdf](https://www.inmujeres.gob.es/disenovovedades/M_MUJERES_Y_DIGITALIZACION_DE_LAS_BRECHAS_A_LOS_ALGORITMOS_04.pdf), consultado 20/04/2022

<sup>16</sup> De similar manera lo apunta, Eguíluz Castañeira, Josu Andoni (2020): "Desafíos y retos que plantean las decisiones automatizadas y los perfilados para los derechos fundamentales", *Estudios de Deusto*, Vol.6, Nº2, p.330.

<sup>17</sup> En este sentido, se debe puntualizar que estos sesgos (indeseados), no son un error del sistema tecnológico o un daño colateral del mismo, sino que son parte y consecuencia de la sociedad en la que vivimos. Precisamente, habitamos en un sistema social basado en la desigualdad, diseñada y estructurada a partir de ella. Por lo tanto, los algoritmos que se alimentan de estos datos que se insertan en este contexto social, irremediamente estarán impregnados de sesgos que reflejan dicha desigualdad.

En un segundo estadio, nos encontramos ante la recolecta y selección de datos que nutrirá el sistema de IA. Esta fase es extremadamente peligrosa en cuanto a la incorporación de sesgos de género, puesto que los datos son un reflejo del rol que la humanidad ha asignado a la mujer en los diferentes ámbitos del quehacer social<sup>18</sup>. Luego, la segunda precaución que debemos adoptar es la selección de los datos, con el objeto de no crear una base de datos sesgada, que reproduzca los estereotipos de género fruto del sexismo imperante de nuestra sociedad<sup>19</sup>. Así, siguiendo con el ejemplo anterior, si seleccionamos como variable la tardanza en denunciar, estaremos creando una base de datos estereotipada, puesto que estaríamos introduciendo una variable que deviene de un perjuicio o estereotipo de género, basado en el comportamiento que debería adoptar una mujer al ser víctima de violencia de género. El algoritmo será entrenado con dicha variable y aprenderá que la tardanza en denunciar es relevante a la hora de valorar la credibilidad del testimonio de la víctima. De esta manera, otorgará un alto nivel de credibilidad a las víctimas que no tardan en denunciar y, en contraposición, un bajo nivel a las que no son capaces, y, atemorizadas, tardan años en denunciar<sup>20</sup>.

Asimismo, además de introducir variables bajo nociones preconcebidas de la realidad, es necesario señalar que se pueden crear bases de datos no representativas, mayoritariamente por infrarrepresentación del género femenino, por la malnutrición de datos históricos que excluyen a las mujeres. Muestra de ello son los conjuntos de datos que utilizan los sistemas de contratación personal en el ámbito empresarial. Estas bases de datos, alimentadas por información histórica sobre contratación, contienen escasas referencias del género femenino. De esta manera, reproducen el estereotipo de la mujer relegada al ámbito privado y doméstico, menoscabando el derecho a la igualdad de trato y de oportunidades de hombres y mujeres en el ámbito laboral<sup>21</sup>.

### *El especial impacto de género de los sistemas de aprendizaje automático y las cajas negras*

El problema de la incorporación de los sesgos o estereotipos en los sistemas de IA, no termina en la creación de bases de datos sesgadas; los nuevos y novedosos algoritmos de autoaprendizaje, el denominado *Machine Learning*, exacerbaba el alcance discriminatorio de estos sistemas. Es determinante entender, que estos algoritmos predicen, pero no explican, y que además pueden encontrar relaciones entre el resultado final que se quiere medir y datos como la edad, el género, la nacionalidad... incluso, emplear datos sin haberlos pedido<sup>22</sup>. Por tanto, el efecto discriminatorio de esta modalidad algorítmica es tan inalcanzable como su naturaleza misma, puesto que, el algoritmo puede desarrollar

---

<sup>18</sup> De similar manera lo expresa, Martínez Garay, Lucía (2018). "Peligrosidad, algoritmos y *due process*: El caso State v. Loomis", *Revista de Derecho Penal y Criminología*, N°20, p.496.

<sup>19</sup> Borges Blázquez, Raquel (2021): "Inteligencia artificial y perspectiva de género: programar, investigar y juzgar con filtro morado", *Revista General de Derecho Procesal*, N°55, p.25.

<sup>20</sup> Este ejemplo lo mencionan, Martínez García, Elena, Borges Blázquez, Raquel, Simó Soler, Elisa (2021): "Inteligencia artificial y perspectiva de género en la justicia penal", *Diario La Ley*, N°47, p.6. Creemos pertinente resaltar que esta base de datos no podría ser legítima en nuestro ordenamiento jurídico, ya que existe una doctrina muy consolidada en cuanto a que la tardanza en denunciar no puede suponer una merma de la credibilidad de la víctima. Hemos optado por este ejemplo para visibilizar el gran alcance y efectos que devienen de la creación de bases de datos sesgadas, en detrimento de los derechos humanos de las mujeres.

<sup>21</sup> Un caso que muestra perfectamente la discriminación algorítmica fundada en bases de datos no representativas lo encontramos en un algoritmo que creó Amazon para utilizarlo en la contratación personal. Los diseñadores del algoritmo emplearon datos de personas que habían prestado funciones en los diez años anteriores en la empresa. La mayoría de estas personas eran hombres, por tanto, la base de datos contenía una sobrerrepresentación del género masculino. De esta manera, el sistema aprendió que ser mujer era equivalente a ser una mala candidata para la empresa, dando lugar a la discriminación algorítmica por razones de género. De manera similar destaca este ejemplo, Soriano Aranz, Alba, Simó Soler, Elisa (2021): "Machine Learning y Derecho: aprendiendo la (des) igualdad", en Barona Vilar, Silvia (Ed.): *Justicia algorítmica y neuroderecho: una mirada multidisciplinaria*, Tirant lo Blanch, Valencia, p.194, al analizar el desarrollo y puesta en funcionamiento de algoritmos discriminatorios.

<sup>22</sup> Pérez Estrada, Miren Josune (2020): "El uso de los algoritmos en el proceso penal y el derecho a un proceso con todas las garantías", en Barona Vilar, Silvia (Ed.): *Claves de la justicia penal: Feminización, Inteligencia artificial, Supranacionalidad y Seguridad*, Tirant lo Blanch, Valencia, p.239.

decisiones discriminatorias sin que podamos identificarla variable que las ha generado.

En consecuencia, la opacidad de este tipo de sistema de IA, tiene un impacto de género particular, dado que la imposibilidad de secuenciar el proceso de la decisión, imposibilita la detección y la posterior corrección de los sesgos de género que impregnan estos sistemas de IA.

## LA INTELIGENCIA ARTIFICIAL Y LA JUSTICIA

### ***La irrupción de la Inteligencia Artificial en la Justicia: de la e-justice al juez robot***

Los sistemas de IA mencionados *supra*, además de modificar sustancialmente aspectos de nuestra vida cotidiana, de igual forma, han generado un cambio de paradigma en el ámbito jurídico. Las nuevas tecnologías ya forman parte de nuestras relaciones jurídicas y paralelamente, se han ido conformando mecanismos de gestión (*e-justice*) y solución de conflictos basados en estas nuevas herramientas “inteligentes” (sistemas de IA predictivos y decisorios). La “justicia algorítmica” es ya una realidad con las oportunidades y riesgos que ello conlleva. Sin subestimar las ventajas y oportunidades que suponen estas nuevas herramientas en los procesos judiciales<sup>23</sup>, nuestra labor se centrará en analizar las amenazas que supone para los derechos fundamentales de las mujeres y, en especial, para el derecho a la igualdad y la no discriminación (art. 14 CE), un uso despersonalizado e insensible al género de estos instrumentos. En esta línea, la reciente propuesta de Reglamento del Parlamento Europeo y del Consejo por el que se establecen las normas armonizadas sobre la IA, publicado el 21 de abril de 2021, ha calificado la aplicación de estos sistemas en el entorno jurídico como de alto riesgo, en virtud de los derechos y garantías que pueden resultar menoscabados<sup>24</sup>.

### ***Modelos de aplicación de los sistemas de IA en la Justicia***

En la aplicación de los sistemas de IA para la resolución de conflictos, se divisan dos modalidades: los sistemas de IA predictivos o asistenciales y los decisorios.

Los primeros se utilizan como elemento de apoyo en la labor jurisdiccional. Es decir, las decisiones que se fundamenten en sistemas de IA se emplean como complemento de la decisión judicial. En el ordenamiento jurídico español, un ejemplo de sistema de IA predictivo o asistencial, lo encontramos en materia de delitos de violencia de género: el Sistema de Seguimiento Integral en Casos de Violencia de Género (en adelante, VioGén)<sup>25</sup>. Concretamente, este modelo predictivo utiliza algoritmos para valorar el riesgo de revictimización, con el objeto de adoptar medidas de protección para la víctima<sup>26</sup>. Es preciso remarcar que este sistema para valorar el riesgo de violencia contra las mujeres es un sistema

---

<sup>23</sup> Entre otras, desde una dimensión procedimental, la IA presenta una utilidad indiscutible para darles solución a los tres grandes problemas que menoscaban la eficiencia de nuestro sistema judicial: la excesiva burocratización, la tardanza en la tramitación, y la escasez de recursos. Para ahondar en las ventajas de la IA en el proceso judicial, véase, Nieva Fenoll, Jordi (2018): *Inteligencia artificial y proceso judicial*, Marcial Pons, Madrid.

<sup>24</sup> La Comisión Europea determina y enumera en qué casos puede la aplicación de la IA suponer un peligro a una autoridad judicial en la investigación, interpretación y aplicación de la Ley. Véase, Comisión Europea (2021): *Proposal for a regulation laying down harmonised rules on artificial intelligence*, Bruselas, COM (2021) 206 final, Anexo III. Documento electrónico, <https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:52021PC0206>, consultado 21/04/2022.

<sup>25</sup> Esta herramienta se creó en el año 2007, en virtud de los mandatos contenidos en los 31.3 y 32 de la Ley Orgánica 1/2004, de 28 de diciembre, de Medidas de Protección Integral contra la Violencia de Género.

<sup>26</sup> González Álvarez, José Luis, López Ossorio, Juan José y Muñoz Rivas, Marina (2018): *La valoración policial del riesgo de violencia contra la mujer pareja en España - Sistema VioGén*, Ministerio del Interior, pp. 18-19.



complementario y asistencial, dado que el agente policial no tiene que aceptar acríticamente el resultado de la evaluación riesgo. El agente puede -y debe- modificar el riesgo apreciado, si aprecia que hay razones que inducen a una estimación errónea del mismo, justificando los motivos de su actuación. De lo contrario, estaría dejando la valoración del riesgo a la mera decisión de un algoritmo, menoscabando el derecho fundamental de las mujeres a vivir en un mundo sin violencia de género.

No obstante, la función decisoria de los sistemas de IA implica sustituir al ser humano y que sea la inteligencia artificial quien juzga el litigio<sup>27</sup>. En este sentido, nos referimos a instrumentos de IA que, no solo permiten complementar la tarea decisoria del juez, sino que se emplean como herramientas sustitutivas de la labor judicial, una suerte de robotización jurisdiccional (juez robot). Esta despersonalización del sistema judicial<sup>28</sup> reviste una vital importancia y trascendencia en el amparo de la igualdad de género, dado que como veremos *infra*, los sistemas de IA pueden convertirse en mecanismos reproductores de desigualdad por los sesgos existentes en los mismos. Por esta razón, es necesario que se prime la labor argumentativa e interpretativa del juez, para poder aplicar el derecho al caso concreto y no adoptar resoluciones que mermen los derechos fundamentales, en general, y los de las mujeres, en particular.

### ***Aplicando los sistemas de IA en el proceso judicial: ¿herramienta productora o garante de la igualdad de género?***

Un ejemplo muy clarificador sobre los riesgos potenciales del uso de los sistemas de IA con funciones decisorias en el proceso judicial penal, en detrimento de los derechos fundamentales de las mujeres, lo encontramos en la sentencia de la Sala Contencioso-Administrativa de la Audiencia Nacional (en adelante, AN) de 30 de septiembre de 2020.

El caso se basa en la deficiente –por no decir nula– protección que otorgaron las Fuerzas de Seguridad del Estado a una mujer que solicitó una orden de protección. Los agentes de la Guardia Civil realizaron a la víctima el cuestionario de cribado del sistema VioGén (VPR), resultando la estimación del riesgo baja. Los agentes, limitándose a la valoración del riesgo apreciado por el sistema VioGén, calificaron el riesgo como “no apreciado”. A pesar de existir indicios para reevaluar dicha estimación –el marido tenía antecedentes de maltrato en su país de origen, la madre del agresor había presenciado episodios de violencia, .... – mediante las actuaciones que los agentes estimasen oportunas y pertinentes, no se modificó y, mediante Auto de 17 de septiembre de 2016, el Juzgado de Primera Instancia e Instrucción nº2 de San Lucar la Mayor, se acordó denegar la orden de protección. Un mes más tarde, la víctima, que la decisión de un algoritmo enmudeció y desprotegió, fue asesinada por su marido.

En este sentido, la AN apreció un funcionamiento negligente de la Guardia Civil “al no ponderar con rigor el riesgo real al que estuvo sometida la víctima, dejándola sin la protección debida, que podía haber evitado la tragedia final”<sup>29</sup>. Por ello, estimando que se daban los requisitos para exigir la responsabilidad patrimonial<sup>30</sup>, condenó al Ministerio del Interior a reparar los daños causados, por el funcionamiento anormal de los servicios policiales.

Este caso evidencia la ausencia de perspectiva de género en la implementación de los sistemas de IA. Por un lado,

---

<sup>27</sup> Martín Diz, Fernando (2021). “Modelos de aplicación de Inteligencia Artificial en Justicia”, en Barona Vilar, Silvia (Ed.): *Justicia algorítmica y neuroderecho: una mirada multidisciplinar*, Tirant lo Blanch, Valencia, p.81

<sup>28</sup> En este sentido, Barona Vilar, Silvia (2019): “Cuarta revolución industrial (4.0) o ciberindustria en el proceso penal: revolución digital, inteligencia artificial y el camino hacia la robotización de la justicia”, *Revista Jurídica Digital UANDES*, Vol.3, Nº1, p.10, habla de una “merma ineludible en la capacidad decisoria del juez, cercenando el razonamiento lógico deductivo desplegado en la formulación de la sentencia”.

<sup>29</sup> SAN 2350/2020, de 30 de septiembre de 2020, FJ nº3. Disponible en: <https://www.poderjudicial.es/search/AN/openDocument/8b1fed1c777df625/20201015>

<sup>30</sup> Dichos requisitos se fijan, entre otras, en la STS 540/2016, de 19 de febrero y en la STS 4942/2014 de 5 de diciembre.

muestra la fe ciega a la objetividad de los sistemas de IA –en este caso VioGén–, olvidando su carácter complementario y asistencial –que no sustitutivo–, depositando, en la valoración de un algoritmo, la toma de decisión que corresponde a la inteligencia humana y al sentido común humano. Por otro lado, pone de manifiesto la ausencia de conocimiento sobre el potencial discriminatorio que entrañan este tipo de sistemas predictivos debido a los sesgos de género que pueden incorporarse, tanto en su diseño, como en la práctica del mismo.

Por tanto, aludimos a la necesidad de juzgar con perspectiva de género y más aún cuando empleamos sistemas de IA, exigiéndose como una obligación jurídica ineludible (art. 4 de la Ley Orgánica 3/2007 para la Igualdad Efectiva de Hombres y Mujeres) para salvaguardar los derechos fundamentales de las mujeres, como lo es en este caso, el derecho a una vida libre sin violencia de género.

## HACIA UNA JUSTICIA E INTELIGENCIA ARTIFICIAL ÉTICA Y GENERO SENSITIVA

Tras analizar de forma somera, los riesgos que entraña la implementación de los sistemas de IA en la Justicia en materia de igualdad y no discriminación, es necesario que dichos riesgos sean atajados de manera urgente, con el objeto de crear sistemas de IA que ayuden (y no obstaculicen) la consecución de la igualdad de género.

### ***Una Europa comprometida con la igualdad de género en el uso de los sistemas de IA***

En Europa existe una gran concienciación de que los derechos fundamentales, en este caso, la igualdad de género, deben ser la columna vertebral de la programación e implementación de los sistemas de IA<sup>31</sup>. En este sentido, la Unión Europea ha desarrollado un marco jurídico que aborda los principios éticos que deben seguirse en el desarrollo de una IA centrada en los seres humanos<sup>32</sup>, que nos sirve de marco de referencia para incluir la perspectiva de género en la IA aplicada en la práctica judicial.

El grupo independiente de Expertos de Alto nivel de Inteligencia Artificial entre los siete requisitos que establece para una IA fiable, en quinto lugar, alude a la diversidad, la no discriminación y la equidad. En esta línea, expresa que, en el contexto de la IA, la igualdad implica que el funcionamiento de este tipo de sistemas no debe generar resultados injustamente sesgados (por ejemplo, los datos utilizados para la formación de los sistemas de IA deben ser lo más inclusivos posibles, de forma que estén representados los diferentes grupos de población). Remarca, que la igualdad, también, requiere un adecuado respeto de las personas y grupos potencialmente vulnerables, mencionando entre otros, a las mujeres<sup>33</sup>.

De similar manera, la propuesta de Reglamento del Parlamento Europeo y del Consejo de 21 de abril de 2021, por el que se establecen las normas armonizadas sobre IA establece la equidad, entre otros, como principio ético y transversal,

---

<sup>31</sup> Un estudio exhaustivo sobre el marco normativo aplicable en esta materia, la encontramos en, Soriano Aranz, Alba (2021): "La aplicación del marco jurídico europeo en materia de igualdad y no discriminación al uso de aplicación de inteligencia artificial", en Valcárcel Fernández, Patricia, Fernández Acevedo, Rafael, Bonorino Ramírez, Pablo Raúl (Coords.): *Nuevas normatividades: Inteligencia Artificial, Derecho y Género*, Thomson Reuters Aranzadi, Pamplona, 63-88.

<sup>32</sup> Para mayor concreción, véase, De Hoyos Sancho, Montserrat (2021): "El libro blanco sobre inteligencia artificial de la Comisión Europea: reflexiones desde las garantías esenciales del proceso penal como sector de riesgo", *Revista Española De Derecho Europeo*, N°76, 9-44.

<sup>33</sup> Comisión Europea (2018): *Directrices éticas para una IA fiable*, Grupo Independiente de Expertos de Alto Nivel sobre Inteligencia Artificial, Bruselas, p. 13 y 38. Documento electrónico, <https://op.europa.eu/es/publication-detail/-/publication/d3988569-0434-11ea-8c1f-01aa75ed71a1>, consultado 21/04/2022.

donde se resalta el compromiso de asegurar una aplicación igualitaria que evite la discriminación directa o indirecta de colectivos determinados como consecuencia de la aparición de sesgos injustos<sup>34</sup>. Asimismo, el Comité Consultivo para la Igualdad de Oportunidades entre Mujeres y Hombres ha elaborado un dictamen que analiza, entre otras, las repercusiones de la IA en la igualdad de género<sup>35</sup>.

### ***La IA al servicio y garantía de la igualdad de género: Medidas para erradicar la discriminación algorítmica***

Consideramos que la implantación de la IA y los algoritmos en la justicia, debe partir de un marco que aúne lo legal y lo ético para evitar amplificar sesgos o prejuicios que vulneran el principio de igualdad y la no discriminación. De esta manera, consideramos esencial que el código informático de los sistemas IA debe adaptarse a los códigos normativos en materia de tutela antidiscriminatoria; la igualdad de género no se negocia con algoritmos.

Dicho esto, en el binomio IA-Justicia, en primer lugar, consideramos esencial crear algoritmos no discriminatorios, equitativos y género-sensitivos<sup>36</sup>. Para ello, los diseñadores de los algoritmos, y, en definitiva, todo el equipo de trabajo que cree el sistema de IA, debe tener formación en materia de género e igualdad (fundamentada en la teoría feminista interseccional)<sup>37</sup>. Asimismo, el grupo de diseñadores debe ser orgánicamente representativo de la diversidad de nuestra sociedad. En definitiva, la perspectiva de género y la interseccionalidad debe estar presente en el diseño, el desarrollo, el despliegue y el uso de los sistemas de IA, especialmente en los que integren el aprendizaje automático (*Machine Learning*)<sup>38</sup>. Asimismo, consideramos de suma importancia regular la opacidad de los algoritmos discriminatorios: su ya demostrada noneutralidad es una amenaza para la garantía de los derechos fundamentales de las mujeres. Por ello, los algoritmos que se quieran emplear en el ámbito jurídico deberán ser auditados y las empresas que los diseñen, deberán contar con guías y planes de igualdad que estudien el impacto de género de los mismos. El Instituto Europeo para la Igualdad de Género (EIGE) y la Agencia de Derechos Fundamentales (FRA) podrían ser los encargados de liderar estas auditorías como órganos independientes y especializados en justicia<sup>39</sup>.

En segundo término, es de capital importancia emplear los sistemas de IA en el proceso judicial, en términos de complementariedad y no de suplantación. Recordemos, las múltiples fases donde los sesgos pueden incorporarse, así como los peligros discriminatorios de los algoritmos de autoaprendizaje y las cajas negras. Es más, debido a que creemos

---

<sup>34</sup> San Miguel Caso, Cristina (2021): "La aplicación de la inteligencia artificial en el proceso: ¿un nuevo reto para las garantías procesales?", *IUS ET SCIENTIA*, Vol.7, N°1, p. 300

<sup>35</sup> Advisory Committee on Equal Opportunities for Women and Men (2020): *AI – opportunities and challenges for gender equality*.

<sup>36</sup> En esta línea, Soriano Aranz, Alba, Simó Soler, Elisa (2021), *Op.cit.*, pp. 200-203, apuntan la posibilidad de utilizar el *Machine Learning* para detectar estereotipos de género en las sentencias.

<sup>37</sup> Los Estados que empleen y las empresas que diseñen los algoritmos, deben tener en consideración las dimensiones interseccionales de la discriminación por razón de género. Las mujeres de ingresos bajos, madres solteras, mujeres de color, mujeres inmigrantes, mujeres con discapacidades y mujeres no heterosexuales pueden verse afectadas de distintas maneras por los sistemas de IA y la automatización. Se deben apreciar las necesidades o expectativas diferenciadas de estos colectivos desde la perspectiva de género interseccional, de lo contrario, no se logrará emplear los sistemas de IA a favor de la igualdad de género.

<sup>38</sup> Algunas de estas medidas fueron propuestas en el Seminario "Justicia algorítmica en clave de género 2021" organizado por el Instituto de las Mujeres y el Ministerio de Igualdad, disponible en: <https://www.youtube.com/watch?v=8x0AMW9ZleM>

<sup>39</sup> Esta posibilidad la recoge, Martínez García, Elena (2021): "Justicia e Inteligencia Artificial sin género", en Barona Vilar, Silvia (Ed.): *Justicia algorítmica y neuroderecho: una mirada multidisciplinar*, Tirant lo Blanch, Valencia, p.222.

extremadamente complicado y utópico crear un sistema carente de sesgos<sup>40</sup>, abogamos por una utilización restrictiva y de carácter asistencial de estas herramientas en la práctica judicial; los sistemas de IA no pueden derivar en la despersonalización del sistema judicial, en detrimento de los derechos fundamentales y de las garantías procesales<sup>41</sup>.

En tercer término, deberá determinarse la posibilidad de crear tribunales especializados en discriminación algorítmica debido a la complejidad y los conocimientos técnicos requeridos para abordar estas cuestiones que se escapan de los instrumentos jurídicos actuales<sup>42</sup>. En esta línea, advertimos que los sistemas de IA no pueden quedarse al margen de la responsabilidad del Estado por no remover los obstáculos que impiden la igualdad real y efectiva de hombres y mujeres (art. 14 y 9.2 CE).

En definitiva, abogamos por la necesidad de que la aplicación de los sistemas de IA en el ámbito jurídico cuente con una *conditio sine qua non*: los sistemas de IA deben emplearse al servicio y garantía del principio de igualdad y la no discriminación<sup>43</sup>. De lo contrario, el avance tecnológico supondría un retroceso en la constante lucha por los derechos fundamentales y libertades de las mujeres.

## BIBLIOGRAFÍA

Advisory Committee on Equal Opportunities for Women and Men (2020): *AI – opportunities and challenges for gender equality*.

Barona Vilar, Silvia (2021): *Algoritmización del derecho y de la justicia. De la Inteligencia Artificial a la Smart Justice*, Tirant lo Blanch, Valencia, 15-728.

Barona Vilar, Silvia (2019): “Cuarta revolución industrial (4.0) o ciberindustria en el proceso penal: revolución digital, inteligencia artificial y el camino hacia la robotización de la justicia”, *Revista Jurídica Digital UANDES*, Vol.3, N°1, 1-21.

Barona Vilar, Silvia (2019): “Inteligencia artificial o la algoritmización de la vida y de la justicia: ¿Solución o problema?”, *Rev. Boliv. de Derecho*, N°20, 18-49.

Borges Blázquez, Raquel (2021): “El sesgo de la máquina en la toma de decisiones en el proceso penal”, *IUS ET SCIENTIA*, Vol.6, N°2, 54-71.

Borges Blázquez, Raquel (2021): “Inteligencia artificial y perspectiva de género: programar, investigar y juzgar con filtro morado”, *Revista General de Derecho Procesal*, N°55, 2-41.

---

<sup>40</sup> En este sentido, Bueno de Mata, Federico (2020): “Macrodatos, inteligencia artificial y proceso: luces y sombras”, *Revista General de Derecho Procesal*, N°51, p.25, remarca que “nunca habrá datos suficientes para garantizar no haya sesgo en tu sistema pues necesitaremos nutrirlo diariamente con millones de datos y patrones que en según qué instancias ni estarán a nuestro alcance”.

<sup>41</sup> San Miguel Caso, Cristina (2021), *Op.cit.*, p.295, remarca que el uso de los algoritmos de autoaprendizaje, así como los sistemas de cajas negras, menoscaban el derecho fundamental a la tutela judicial efectiva (art.24 CE), al situar al perjudicado en una situación de indefensión debido al desconocimiento de las causas que han generado el trato discriminatorio.

<sup>42</sup> Así lo apunta, Martín Ríos, Pilar (2020): “Problemas de admisibilidad de la prueba obtenida de dispositivos de almacenamiento digital”, *Revista General de Derecho Procesal*, N°51, p.6. De similar manera, lo manifiesta Eguiluz, Castañeira, Josu Andoni (2020), *Op.cit.*, p.360.

<sup>43</sup> En esta línea, Barona Vilar, Silvia (2019): “Inteligencia artificial o la algoritmización de la vida y de la justicia:

¿Solución o problema?”, *Rev. Boliv. de Derecho*, N°20, p.46 remarca que “La Inteligencia Artificial es un buen instrumento siempre que sea un instrumento que sirva para conseguir un mundo mejor, un mundo en el que se erradique la desigualdad, la pobreza y la falta de oportunidades. De lo contrario, la máquina inteligente destruirá el mundo de los humanos”.

Bueno de Mata, Federico (2020): "Macrodatos, inteligencia artificial y proceso: luces y sombras", *Revista General de Derecho Procesal*, N°51, 1-31.

Comisario de Derechos Humanos del Consejo de Europa (2019): *Unboxing Artificial Intelligence: 10 steps to protect Human Rights*, 5-24. Documento electrónico, <https://rm.coe.int/unboxing-artificial-intelligence-10-steps-to-protect-human-rights-reco/1680946e64>, consultado: 19/04/2022.

Comisión Europea (2019). *A definition of AI: Main capabilities and scientific disciplines*, Bruselas. Documento electrónico, <https://www.aepd.es/sites/default/files/2019-12/ai-definition.pdf>, consultado: 18/04/2022.

Comisión Europea (2018): *Directrices éticas para una IA fiable*, Grupo Independiente de Expertos de Alto Nivel sobre Inteligencia Artificial, Bruselas. Documento electrónico, <https://op.europa.eu/es/publication-detail/-/publication/d3988569-0434-11ea-8c1f-01aa75ed71a1>, consultado 21/04/2022.

Comisión Europea (2021): *Proposal for a regulation laying down harmonised rules on artificial intelligence*, Bruselas. Documento electrónico, <https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:52021PC0206>, consultado 21/04/2022.

De Hoyos Sancho, Montserrat (2021): "El libro blanco sobre inteligencia artificial de la Comisión Europea: reflexiones desde las garantías esenciales del proceso penal como sector de riesgo", *Revista Española De Derecho Europeo*, N°76, 9-44.

Eguiluz, Castañeira, Josu Andoni (2020): "Desafíos y retos que plantean las decisiones automatizadas y los perfiles para los derechos fundamentales", *Estudios de Deusto*, Vol.6, N°2, 325-367.

Martínez García, Elena, Borges Blázquez, Raquel, Simó Soler, Elisa (2021): "Inteligencia artificial y perspectiva de género en la justicia penal", *Diario La Ley (Ciberderecho)*, N°47.

Facio, Alda, Fries, Lorena (1999): *Género y Derecho*, LOM ediciones, Santiago de Chile.

González Álvarez, José Luis, López Ossorio, Juan José y Muñoz Rivas, Marina (2018): *La valoración policial del riesgo de violencia contra la mujer pareja en España - Sistema VioGén*, Ministerio del Interior, Madrid, 2-120.

Instituto de la Mujer (2020): *Mujeres y digitalización. De las brechas a los algoritmos*, Catálogo de Publicaciones de la Administración General del Estado, Madrid, 3-126. Documento electrónico, [https://www.inmujeres.gob.es/disenov/novedades/M\\_MUJERES\\_Y\\_DIGITALIZACION\\_DE\\_LAS\\_BRECHAS\\_A\\_LOS\\_ALGORITMOS\\_04.pdf](https://www.inmujeres.gob.es/disenov/novedades/M_MUJERES_Y_DIGITALIZACION_DE_LAS_BRECHAS_A_LOS_ALGORITMOS_04.pdf), consultado 20/04/2022.

Jesús Ignacio (2018): "Derecho Inteligente", *Cuadernos Electrónicos de filosofía del Derecho*, N°37, 95-114.

Martín Ríos, Pilar (2020): "Problemas de admisibilidad de la prueba obtenida de dispositivos de almacenamiento digital", *Revista General de Derecho Procesal*, N°51, 2-31.

Martínez García, Elena (2021): "Justicia e Inteligencia Artificial sin género", en Barona Vilar, Silvia (Ed.): *Justicia algorítmica y neuroderecho: una mirada multidisciplinar*, Tirant lo Blanch, Valencia, 21-712.

Martínez Garay, Lucía (2018). "Peligrosidad, algoritmos y due process: El caso State v. Loomis", *Revista de Derecho Penal y Criminología*, N°20, 485-502.

Nieva Fenoll, Jordi (2018): *Inteligencia artificial y proceso judicial*, Marcial Pons, Madrid.

O'Donnell, Renata M. (2019): "Challenging racist predictive policing algorithms under the equal protection clause", *New York University Law Review*, Vol. 9, N°3, 544-580.

Pérez Estrada, Miren Josune (2020): "El uso de los algoritmos en el proceso penal y el derecho a un proceso con todas las garantías", en Barona Vilar, Silvia (Ed.): *Claves de la justicia penal: Feminización, Inteligencia artificial, Supranacionalidad y Seguridad*, Tirant lo Blanch, Valencia, 235-253.

San Miguel Caso, Cristina (2021): "La aplicación de la inteligencia artificial en el proceso: ¿un nuevo reto para las garantías procesales?", *IUS ET SCIENTIA*, Vol.7, N°1, 286-303.

Solar Cayón, José Ignacio (2019): *La inteligencia artificial jurídica. El impacto de la innovación tecnológica en la práctica del Derecho y el mercado de servicios jurídicos*, Aranzadi, Pamplona

Soriano Aranz, Alba (2021): "La aplicación del marco jurídico europeo en materia de igualdad y no discriminación al uso

de aplicación de inteligencia artificial”, en Valcárcel Fernández, Patricia, Fernández Acevedo, Rafael, Bonorino Ramírez, Pablo Raúl (Coords.): *Nuevas normatividades: Inteligencia Artificial, Derecho y Género*, Thomson Reuters Aranzadi, Pamplona, 63-88.

Soriano Aranz, Alba, Simó Soler, Elisa (2021): “Machine Learning y Derecho: aprendiendo la(des) igualdad”, en Barona Vilar, Silvia (Ed.): *Justicia algorítmica y neuroderecho: una mirada multidisciplinar*, Tirant lo Blanch, Valencia, 183-204.

Schwab, Klaus (2017): *The Fourth Industrial Revolution*, Crown Business Publishing, New York.

World Economic Forum (2021): *Global Gender Gap Report 2021*, Suiza, 2-405. Documento electrónico, <https://www.weforum.org/reports/global-gender-gap-report-2021>, consultado 20/04/2022.