

Género y herramientas algorítmicas: una revisión sistemática sobre los retos y el impacto que plantean en el ámbito judicial y penitenciario

Rocío Martínez Almanza

Centro CRÍMINA, Universidad Miguel Hernández de Elche

Sandra Pérez Domínguez

Centro CRÍMINA, Universidad Miguel Hernández de Elche

Fecha de presentación: octubre 2025

Fecha de aceptación: marzo 2026

Fecha de publicación: octubre 2026

Resumen

En los últimos años, el uso de herramientas algorítmicas de valoración del riesgo se ha extendido en los sistemas judicial y penitenciario, con la promesa de aportar mayor objetividad y eficiencia en la toma de decisiones. Sin embargo, su incorporación genera un creciente debate en torno a los posibles sesgos y desigualdades que pueden reproducir, especialmente en lo que respecta al género. El presente estudio realiza una revisión sistemática de la literatura sobre el impacto del género en el uso de herramientas algorítmicas de valoración del riesgo en los sistemas judicial y penitenciario, con la finalidad de identificar si el diseño y los datos de entrenamiento de estas tecnologías reproducen o amplifican las desigualdades de género. Para ello, se analizaron estudios publicados entre 2007 y 2024, seleccionados conforme al protocolo PRISMA 2020, a partir de bases de datos especializadas y literatura complementaria. Los resultados, basados en siete trabajos relevantes, evidencian que la aplicación de algoritmos predictivos puede derivar en decisiones más severas hacia las mujeres, la estigmatización, la patologización de conductas femeninas y las vulneraciones de derechos fundamentales. También se observa la infrarrepresentación de ciertos colectivos y el riesgo de exclusión social. Entre las estrategias propuestas para mitigar estos impactos destacan las auditorías periódicas, la inclusión de la perspectiva de género desde el diseño, la supervisión humana en la toma de decisiones y la formación específica de los operadores judiciales. Los hallazgos subrayan la necesidad de un desarrollo y uso de estas herramientas que priorice la equidad y transparencia.

Palabras clave

impacto de género; herramientas algorítmicas; sistema judicial; sistema penitenciario; inteligencia artificial

Gender and algorithmic tools: a systematic review on the challenges and impact they pose in the judicial and prison systems

Abstract

In recent years, the use of algorithmic tools for risk assessment has expanded within judicial and prison systems, promising greater objectivity and efficiency in decision-making. However, their implementation has sparked debate about potential biases and inequalities they may reproduce, particularly with respect to gender. This study conducts a systematic review of the literature on the impact of gender on the use of algorithmic tools in judicial and correctional systems, aiming to determine whether the design and training data of these technologies reproduce or amplify gender inequalities. Studies published between 2007 and 2024 were analyzed, selected according to PRISMA 2020. Results, based on seven relevant studies, indicate that predictive algorithms can lead to harsher decisions against women, stigmatization, pathologization of female behaviours, and violations of fundamental rights. Underrepresentation of certain groups and the risk of social exclusion were also observed. Strategies to mitigate these impacts include regular audits, the incorporation of a gender perspective from the design stage, human oversight in decision-making, and targeted training for judicial operators. The findings underscore the need to develop and use these algorithmic tools prioritizing equity and transparency.

Keywords

gender impact, algorithmic tools, justice system, penitentiary system, artificial intelligence

Introducción

La transformación tecnológica, impulsada por las herramientas digitales y de inteligencia artificial (IA) en el sistema de justicia penal, genera un debate en torno a las potencialidades y los riesgos que conlleva. Algunos autores defienden que esta transformación trae consigo una mejora de la eficiencia y eficacia en el trabajo desarrollado en este ámbito, así como la precisión en las tareas legales que se lleven a cabo (Farfán *et al.*, 2023; Montesinos, 2024). Casanovas, Rodríguez-Doncel y González-Conejero (2017) señalaron que «la inteligencia artificial está transformando rápidamente el campo del Derecho, desafiando las nociones tradicionales de responsabilidad, privacidad y equidad en el sistema legal». No obstante, también se ha evidenciado que su uso puede conllevar riesgos que afectan a derechos fundamentales (Farfán *et al.*, 2023), que pueden quedar ocultos bajo prácticas que aparentan una neutralidad tecnológica (Wachter-Boettcher, 2017). Esta situación se ve agravada por la opacidad algorítmica, entendida como la falta de transparencia sobre cómo los sistemas llegan a sus decisiones, lo que dificulta su supervisión y el control necesario para evitar vulneraciones en los derechos fundamentales (Pasquale, 2015). Ante este

panorama, diversos organismos internacionales han señalado la importancia de desarrollar e incorporar marcos normativos que controlen el desarrollo y uso que se le da a la IA, basados en derechos humanos, equidad y transparencia (Consejo Europeo, 2020; UNESCO, 2020).

Estas recomendaciones evidencian que el desafío no se limita únicamente al plano de la implementación, sino que se exige un replanteamiento de las garantías sociales, jurídicas y éticas que esta transformación plantea. En este sentido, resulta imprescindible que la transformación tecnológica vaya acompañada de la incorporación de un enfoque de género que garantice el respeto al principio de igualdad consagrado en nuestra Constitución (Barona, 2019), el cual podría verse afectado por sesgos que podrían propiciar situaciones de discriminación (Harcourt, 2015; Starr, 2014). Aunque el impacto del género en la toma de decisiones judiciales no es inherente a la IA, este suele estar condicionado por las decisiones previas en la construcción de las bases de datos que alimentan las herramientas algorítmicas, así como por la forma en la que se diseñan los algoritmos predictivos (Bustelo, 2025; Buolamwini y Gebru, 2018; Corvalan, *et al.*, 2021; Saavedra-Vera, Jáuregui-Bustamante y Arista-Bustamante, 2023).

Estas decisiones, lejos de ser neutras, pueden reproducir y amplificar sesgos existentes en el sistema judicial y policial, lo que perpetúa desigualdades de género que afectan significativamente a las mujeres y otros grupos marginados (Xenidis y Senden, 2020). Como ya planteaba Mayson (2018) con la expresión «bias in, bias out», los sesgos que reproducen y amplifican los algoritmos predictivos son el resultado de sesgos humanos previamente incorporados.

En los últimos años, el análisis del sesgo algorítmico ha despertado un notable interés académico, lo que ha generado una amplia producción que evidencia la influencia de variables sociodemográficas como la etnia, la edad o el género en los resultados de los sistemas automatizados. Un ejemplo de ello es el estudio de Sarrius *et al.* (2023), en el que se analizaron sistemas de *face verification* y se demostró que los algoritmos de reconocimiento facial presentan disparidades significativas en función de la raza, la edad y el género. En particular, grupos como las mujeres africanas mayores registran tasas de error considerablemente más altas que otros colectivos, lo que evidencia cómo múltiples variables sociodemográficas pueden afectar el desempeño de los sistemas automatizados. Asimismo, An *et al.* (2025) evaluaron modelos de lenguaje (LLM) utilizados para puntuar currículos laborales y encontraron que estos algoritmos favorecen sistemáticamente a determinados grupos sobre otros, generando impactos diferentes según el género y la raza, incluso cuando las habilidades son equivalentes.

En este marco de discusión, y dada la amplitud del debate, la presente revisión se centró específicamente en el impacto del género en el uso de herramientas algorítmicas de valoración del riesgo en los ámbitos judicial y penitenciario. Esta elección responde a la creciente relevancia de examinar cómo las diferencias de género pueden incidir en la toma de decisiones mediadas por la IA, especialmente en un contexto en el que estas tecnologías adquieren un papel cada vez más influyente en las cuestiones vinculadas a la libertad y los derechos fundamentales de las personas.

Con este propósito, consideramos pertinente organizar y sistematizar la evidencia disponible en la literatura académica sobre el uso de herramientas algorítmicas en los sistemas judicial y penitenciario, para ofrecer un análisis riguroso que contribuya a enriquecer la comprensión del papel que desempeña el género en la interacción entre IA, toma de decisiones automatizada y derechos fundamenta-

les. El presente estudio tiene como objetivo analizar en profundidad el impacto del género en la literatura académica vinculada al uso de herramientas algorítmicas en el ámbito judicial y penitenciario. Para ello, se llevó a cabo una revisión sistemática con la finalidad de identificar los diferentes estudios que abordan tanto los impactos de género como las estrategias de mitigación propuestas en este contexto. Mediante esta revisión, se pretende proporcionar una visión comprensiva sobre las implicaciones sociales, éticas y jurídicas de la implementación de la IA en la justicia, y destacar las medidas que pueden contribuir a una mayor equidad en materia de género y transparencia en los procesos judiciales y penitenciarios. Se espera que los resultados de esta revisión no solo ofrezcan una síntesis crítica de los estudios existentes, sino que también proporcionen recomendaciones sobre cómo los sistemas algorítmicos pueden ser diseñados y utilizados de manera que respeten los principios fundamentales de igualdad y no discriminación en general y en materia de género en particular.

1. Método

1.1. Procedimiento

Con la finalidad de responder a la pregunta de investigación relativa a la posible existencia de un impacto de género en el uso de herramientas algorítmicas de valoración del riesgo en el sistema de justicia penal y penitenciario, se optó por el uso de la revisión sistemática de la literatura como metodología de investigación. Esta elección se fundamenta en su capacidad para sintetizar de manera rigurosa la información científica disponible en relación con el objetivo planteado, así como para identificar posibles vacíos en la evidencia existente, lo que facilita la delimitación de áreas que requieren investigación adicional (Ferreira *et al.*, 2011). Asimismo, este enfoque metodológico permite identificar, evaluar de forma crítica y sintetizar de manera ordenada la evidencia científica relevante en relación con una pregunta de investigación previamente definida (Centro Cochrane Iberoamericano, 2012), exigiendo la aplicación de procedimientos rigurosos que garanticen una adecuada identificación y evaluación de los resultados obtenidos (Centeno *et al.*, 2021). En términos generales, la revisión sistemática puede entenderse como un sistema orientado a localizar la información relevante, evaluar su calidad y sintetizar los resultados disponibles (Salcido *et al.*, 2021).

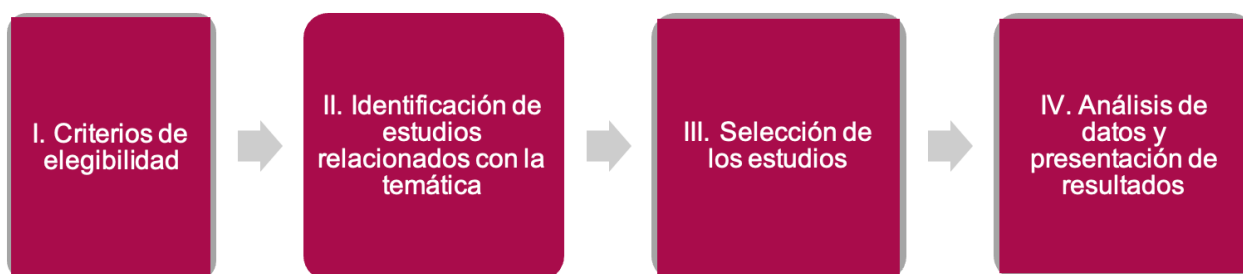
El proceso de revisión se desarrolló siguiendo los criterios de calidad establecidos en el protocolo PRISMA 2020 (Page *et al.*, 2021). Del mismo modo, con el objetivo de garantizar la transparencia, exhaustividad y reproducibilidad de cada una de las etapas de la revisión sistemática, se utilizó la herramienta PRISMA Checklist (PRISMA, n.d.)¹ que se incluye en el anexo.²

Para el desarrollo de las distintas fases del proceso se tomaron como referencia revisiones sistemáticas previas, adaptadas desde el ámbito de la salud al de las ciencias sociales y jurídicas (Cooper & Hedges, 1994; Ferreira *et*

al., 2011; Centro Cochrane Iberoamericano, 2011; Moher *et al.*, 2009; Pardal-Refoyo y Pardal-Peláez, 2020; Pérez-Domínguez, 2024; Pérez-Domínguez, Castro-Toledo y Miró-Llinares, 2019; Reina, 2019). Estos trabajos sirvieron como guía para estructurar las principales etapas del procedimiento, incluyendo la formulación de los objetivos, la definición de la estrategia de búsqueda (y de las palabras clave), la selección de los estudios de interés y su posterior análisis.

A partir de lo anterior, el desarrollo de la presente revisión sistemática se llevó a cabo tal como muestra la figura 1.

Figura 1. Procedimiento de la revisión sistemática



1.2. Criterios de elegibilidad

Siguiendo el objetivo del estudio, identificar el impacto de género en el uso de herramientas algorítmicas de valoración del riesgo, en la presente revisión se incluyeron estudios que abordan diferentes aspectos claves relacionados con el objetivo de la investigación. Con ello, la búsqueda se realizó tanto a escala nacional como internacional, considerando la posibilidad de posibles impactos que no se contemplen en el uso de herramientas utilizadas en el ámbito nacional. La selección de los artículos que finalmente fueron analizados en esta revisión se realizó siguiendo una serie de criterios de inclusión y exclusión, entre los que destacan los siguientes:

- **Criterio 1:** se incluyeron únicamente artículos publicados en revistas científicas y tesis doctorales disponibles en acceso abierto. La elección de fuentes revisadas por pares responde a la necesidad de asegurar la calidad y

validez científica de los trabajos incluidos. Asimismo, el requisito de acceso abierto busca garantizar la transparencia del proceso y la posibilidad de replicación del estudio por parte de otros investigadores.

- **Criterio 2:** se estableció como filtro el idioma, seleccionando exclusivamente publicaciones en inglés y español. Aunque se aplicó esta restricción lingüística, no se identificaron estudios relevantes en otros idiomas durante la búsqueda.
- **Criterio 3:** se incluyeron únicamente estudios que abordaran explícitamente la dimensión de género en relación con el uso de herramientas algorítmicas de valoración del riesgo. Este criterio respondía a la finalidad central del estudio de examinar cómo los sesgos y las desigualdades de género pueden verse reproducidos o mitigados en la aplicación de dichas herramientas en contextos judiciales y penitenciarios.

1. La lista de verificación en formato web puede cumplimentarse a través del siguiente enlace: <https://prisma.shinyapps.io/checklist/>

2. Se puede acceder al mismo por medio de este link: https://drive.google.com/file/d/13BNw-if2yPABvF-QKSGmMjIBInUse52U/view?usp=share_link

1.3. Identificación de estudios relacionados con la temática

En primer lugar, se realizó una revisión exploratoria de la literatura (*scoping review*), con la finalidad de mapear el panorama académico en torno al tema propuesto. Esto permitió identificar qué palabras clave debían tomarse en consideración en el momento de realizar la búsqueda de la literatura, así como optimizar el proceso de búsqueda en las diferentes bases de datos. La búsqueda de estudios se llevó a cabo en marzo de 2025, mediante las bases de datos ProQuest, Scopus y Web of Science, plataformas seleccionadas por su alto nivel de confiabilidad y por su amplio alcance en diversas disciplinas académicas. Por otro lado, y como resultado de la *scoping review* realizada en un primer momento, se incluyeron una serie de artículos que, coincidiendo con los criterios de inclusión y exclusión señalados, se identificaron como relevantes para la revisión sistemática, que no se localizaron en las bases de datos revisadas posteriormente.

Cabe mencionar que la selección de la literatura se realizó de manera individual por parte de las dos investigadoras; posteriormente, se pusieron en tela de juicio aquellas en las que había discrepancia y una tercera persona externa actuó de juez para determinar si el artículo era apto o no para la revisión.

1.4. Búsqueda de la literatura

La búsqueda bibliográfica, como ya se ha mencionado, se llevó a cabo en tres bases de datos principales: ProQuest, Scopus y Web of Science, seleccionadas por su relevancia en los campos de las ciencias sociales, el derecho y la criminología.

Se utilizaron las siguientes palabras clave:

- **ProQuest:** (impact) AND (Gender) AND ('Algorithmic tools' OR 'risk assessment') AND ('justice' OR 'judicial' OR 'prisons' OR 'Prisons' OR 'Penitentiary')

- **Scopus:** TITLE-ABS-KEY(("impact") AND ("Gender") AND ("Algorithmic tools" OR "risk assessment")) AND (("justice" OR "judicial" OR "prisons" OR "Penitentiary"))
- **Web of Science:** ("impact") AND ("Gender") AND ("Algorithmic tools" OR "risk assessment") AND (("justice" OR "judicial" OR "prisons" OR "Penitentiary"))

La razón por la cual existen diferencias en la manera en la que se presentan las palabras clave utilizadas se debe a los criterios que emplea cada base de datos para realizar las búsquedas.

1.5. Selección de la literatura

Una vez aplicados todos los criterios de inclusión y exclusión, así como realizada la revisión para descartar aquellos posibles artículos duplicados por medio de la aplicación Rayyan, se identificó un total de 58 artículos potenciales. A partir de ahí, se procedió a la revisión del *abstract* de cada uno de ellos, con la finalidad de determinar si realmente se vinculan con el objetivo de la revisión, lo que permitió descartar 51 artículos.

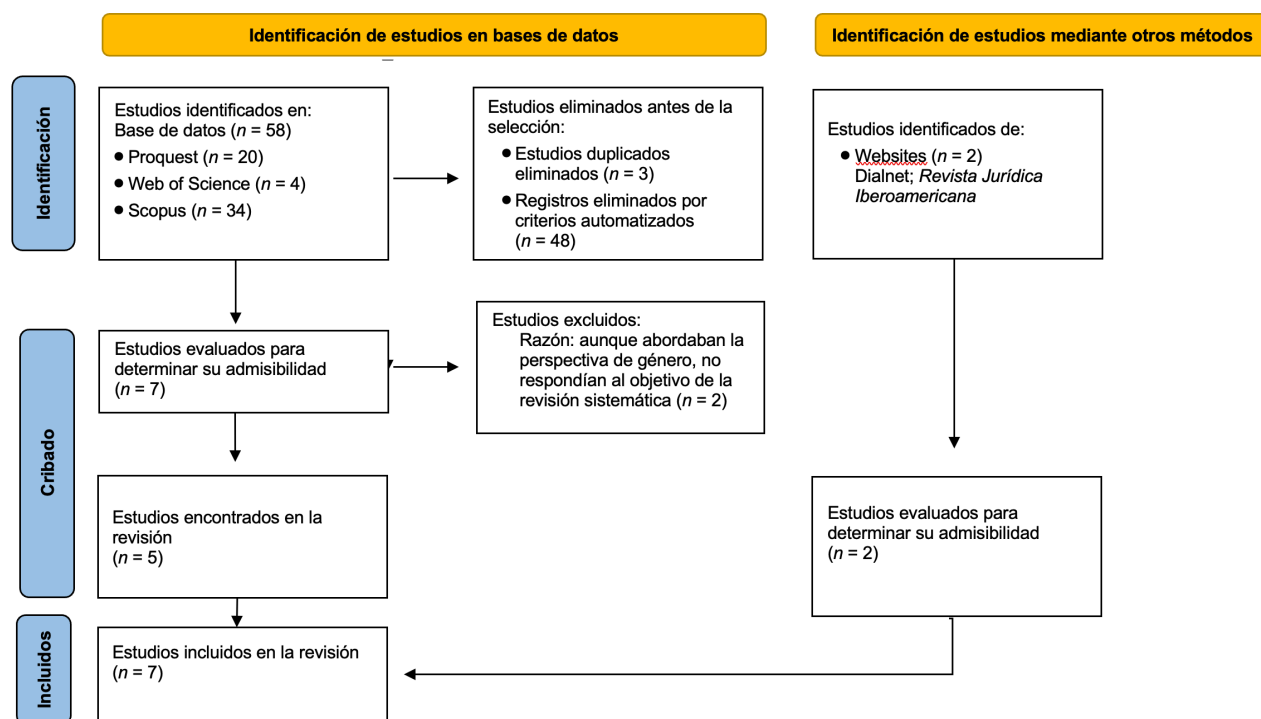
Una vez realizada esta primera revisión, se llevó a cabo una revisión en profundidad, lo que dio lugar a una selección final de siete artículos que se corresponden con los criterios establecidos, contemplando en dicha cifra los dos artículos hallados fuera de la base de datos.

Durante el proceso de revisión, se identificaron dos artículos adicionales que no cumplían estrictamente con los criterios de inclusión predefinidos, como el estar en alguna de las bases de datos revisadas. Con el fin de aportar un marco más completo para la interpretación de los resultados, se decidió incluirlos en el análisis. Estos artículos procedían de Dialnet y de la *Revista Actualidad Jurídica Iberoamericana*.³

En la figura 2, se presenta el diagrama de flujo que resume las etapas del proceso seguido en la revisión sistemática.

3. Los artículos se pueden revisar en los siguientes enlaces: <https://dialnet.unirioja.es/servlet/articulo?codigo=9725881> y https://revista-aji.com/wp-content/uploads/2024/07/AJ121_Art20.pdf

Figura 2. Estudios identificados en las bases de datos



1.6. Extracción de datos y análisis

La extracción de datos se organizó en una tabla resumen (anexo 1),⁴ estructurada según las variables necesarias para el cumplimiento del objetivo del estudio, con la finalidad de facilitar el análisis. En concreto, se recopilaban las siguientes variables: autor(es), año de publicación, objetivos, método, herramienta que se analiza (contexto y población), impacto identificado y recomendaciones/propuestas para mitigar dichos impactos.

En relación con la variable «herramienta que se analiza», se consideró como tal a la herramienta que los estudios ponen en evaluación, incluyendo su contexto de uso y la población a la que se aplica. Por su parte, «impacto identificado» hace referencia a los efectos que los estudios señalan como vinculados al género y que pueden generar consecuencias positivas o negativas en los procesos de decisión judicial derivados del uso de algoritmos de va-

loración del riesgo. Finalmente, la variable «recomendaciones/propuestas para mitigar dichos impactos» recoge las soluciones planteadas en los distintos estudios para reducir la influencia del género en la utilización de herramientas algorítmicas de valoración del riesgo.

1.7. Evaluación de la calidad de los estudios seleccionados

Con la finalidad de garantizar que los estudios incluidos en esta revisión sistemática alcanzaran unos estándares mínimos de calidad, se empleó el método de evaluación propuesto por el Critical Appraisal Skills Program (CASP). Esta herramienta facilita el análisis del rigor metodológico, la consistencia de los resultados y la relevancia de cada investigación (Long *et al.*, 2020). Los resultados de la aplicación de esta evaluación a las publicaciones seleccionadas se recogen en una tabla, ubicada en el anexo 2.⁵ En esta tabla, un ✓ (sí) indica que el criterio se cumple en

4. El anexo 1 está disponible en: https://drive.google.com/file/d/1_vSoJcHBDiQOWInishxY0YMOGeeQsAW/view?usp=share_link

5. El anexo 2 está disponible en: https://drive.google.com/file/d/173OrLDamit7bfCxOdoU0mnEmraOFDBm3/view?usp=share_link

su totalidad; un $\pm$ (parcialmente) refleja que se satisfacen algunos aspectos, pero persisten limitaciones; y una $\times$ (no) señala que el criterio no se cumple. Cabe destacar que todos los estudios evaluados fueron finalmente incluidos en la revisión, por lo que ninguno de los artículos que había superado la fase de cribado inicial fue descartado en esta etapa del análisis.

2. Resultados

Tras el proceso de cribado, se realizó una revisión en profundidad de los artículos seleccionados ($n = 7$). De los cuales cuatro ($n = 4$) se encontraban en el idioma español y tres ($n = 3$) en idioma inglés. La temporalidad de los artículos va del año 2007 al año 2024.

Desde una perspectiva metodológica, los estudios incluidos en esta revisión sistemática presentan una orientación predominantemente cualitativa, con una tendencia al uso de análisis de casos, entrevistas a personas expertas y actores institucionales, revisión documental y análisis del discurso para examinar el funcionamiento y las implicaciones sociales de las herramientas algorítmicas de valoración del riesgo. Esto ha permitido profundizar en la comprensión del impacto de género asociado a su aplicación en el ámbito penal. Sin embargo, solo algunos trabajos incorporan análisis cuantitativos o técnicas estadísticas orientadas a la detección de patrones de sesgo o disparidad, y aún menos aplican modelos avanzados como regresiones logísticas o algoritmos de clasificación, lo que plantea interrogantes sobre la solidez empírica de ciertas conclusiones, muchas veces sustentadas en marcos interpretativos y debates teóricos.

Tabla 1. Resumen de las características de los estudios

Categoría	Tipo de estudio	n	%
Tipo de estudio	Cualitativo	5	71,40 %
	Cuantitativo	2	28,60 %
Idioma	Español	4	57,14 %
	Inglés	3	42,86 %

Los estudios revisados pueden agruparse en tres categorías principales. La mayoría analiza los riesgos de la IA y de los algoritmos predictivos en contextos judiciales y de violencia de género ($n = 5$), destacando que el diseño y entrenamiento de estos sistemas sin perspectiva de género y con datos sesgados tiende a reproducir y amplificar las desigualdades estructurales, afectando derechos fundamentales, estigmatizando colectivos y perpetuando prejuicios. Un segundo grupo ($n = 2$) examina el impacto de los algoritmos en la valoración del riesgo y la prisión preventiva, identificando disparidades de género y raza y la necesidad de ajustes en las herramientas junto con un mayor control humano para mitigar los errores. Finalmente, un tercer grupo ($n = 1$) se enfoca en la justicia juvenil, mostrando cómo los sistemas de evaluación de riesgo patologizan la conducta de las chicas, especialmente de jóvenes racializadas, y refuerzan las desigualdades bajo una apariencia de objetividad técnica. La mayoría de los estudios combina revisiones bibliográficas y análisis de casos con entrevistas a expertos, y solo de forma minoritaria análisis estadísticos mediante *machine learning* para evaluar los sesgos y las disparidades en la toma de decisiones. En el anexo 1^o se presentan con mayor detalle los resultados encontrados.

Ahora bien, teniendo en consideración que el principal objetivo de esta revisión fue identificar los elementos que la literatura ha señalado como posibles generadores de impacto de género en el uso de algoritmos predictivos dentro del sistema judicial y penitenciario, los resultados se organizaron en amplias categorías analíticas que sistematizan y facilitan la interpretación de los hallazgos. En la tabla 2 se recogen estos impactos, agrupados por temática.

En la revisión de los artículos seleccionados se identificaron varios aspectos claves relacionados con el uso de herramientas algorítmicas en el ámbito judicial y el impacto de género, especialmente en contextos de violencia de género. Los estudios seleccionados proporcionan una visión comprensiva, abordando tanto los impactos negativos como los positivos y las recomendaciones para mitigar sus efectos adversos.

6. El anexo 1 está disponible en: https://drive.google.com/file/d/1_vSoJcHBDiQOWInishxYOYMOGeeQsAW/view?usp=share_link

Tabla 2. Resumen de los impactos asociados a los sesgos de género en el uso de algoritmos predictivos

Categoría	Impacto	N.º de trabajos
Desigualdad y discriminación estructural	Refuerza desigualdades de género y contribuye a la discriminación estructural.	3
	Reproduce sesgos algorítmicos que afectan negativamente a los colectivos infrarrepresentados.	2
	Genera exclusión social, especialmente hacia las mujeres.	4
Violencia de género y estigmatización	Perpetúa desigualdades que favorecen la violencia de género.	2
Sesgos en datos y algoritmos	Aplica medidas judiciales más severas a las mujeres, como la prisión preventiva.	2
	Tiende a patologizar las conductas de mujeres jóvenes, reforzando estigmas.	1
Derechos fundamentales y garantías procesales	Riesgo de vulneración de las garantías procesales y los derechos fundamentales, especialmente los vinculados a la igualdad y a una vida libre de violencia.	2
Mejora en las valoraciones	La herramienta puede contribuir a reducir la disparidad de género frente a decisiones judiciales tradicionales.	1

Los resultados de la revisión muestran que las herramientas algorítmicas utilizadas en el ámbito judicial y penitenciario tienden a reproducir y, en algunos casos, amplificar las desigualdades estructurales de género. En particular, se observó que los algoritmos analizan datos sesgados, lo que puede resultar en decisiones discriminatorias hacia las mujeres y otros colectivos vulnerables. Por ejemplo, en los estudios de Montesinos (2023) y López (2024), se destacó cómo la incorporación de sesgos de género en los datos de entrenamiento puede perpetuar desigualdades, especialmente cuando los sistemas no están diseñados para reconocer ni corregir estos sesgos. De manera similar, Barron (2007) y Faddoul *et al.* (2020) identificaron que las herramientas de evaluación de riesgo, como la Public Safety Assessment (PSA), pueden generar decisiones más severas para las mujeres en el ámbito de la prisión preventiva, y en el contexto de la justicia juvenil, las herramientas tienden a patologizar las conductas de las mujeres jóvenes, especialmente las racializadas, al considerar ciertos comportamientos como señales de riesgo elevado.

Además, se encontró que la utilización de la variable de género para predecir comportamientos delictivos puede generar escenarios de marginación para colectivos sociales ya vulnerables. Llorente (2022) argumentó que esto puede contribuir a la deshumanización del proceso judicial, al priorizar los algoritmos sobre la interpretación contextual y humana de cada caso.

En la tabla 3 se resumen las principales estrategias identificadas que podrían mitigar los impactos mencionados.

Tabla 3. Resumen de medidas sugeridas para mitigar el impacto

Medidas
Algoritmos auditados y controlados.
Exigencias de responsabilidad ante decisiones tomadas por la IA.
No sustituir decisiones humanas por herramientas automáticas.
Detectar y neutralizar los sesgos algorítmicos (con medidas políticas, jurídicas y tecnológicas).
Evitar la incorporación de sesgos de género y estereotipos discriminatorios.
Incorporar la perspectiva de género en el diseño, desarrollo y uso en justicia.
Equipos formados en género e igualdad para el diseño del sistema.
Propuesta del modelo Handoff Tree para mitigar los impactos predictivos y dejar la decisión final al juez.
Formulación explícita de la imparcialidad frente a los prejuicios raciales y de género.
Ajuste de PSA para incluir raza y género donde sea legal.
Crítica a la ausencia de medidas de mitigación de impactos.
Reconocimiento y corrección de sesgos algorítmicos inherentes.
Formación crítica a operadores de justicia para evitar las lecturas automáticas de los resultados algorítmicos.

Con la finalidad de complementar los resultados en relación con el impacto de género en el uso de herramientas algorítmicas de valoración del riesgo, también se identificaron varias condiciones necesarias para el uso justo y equitativo de estas herramientas en el ámbito penitenciario y judicial. Un aspecto comúnmente señalado fue la necesidad de incorporar una perspectiva de género en el diseño, desarrollo y aplicación de las herramientas algorítmicas. Montesinos (2023) y López (2024) sugieren que, para evitar la perpetuación de sesgos, los algoritmos deben ser desarrollados por equipos interdisciplinarios que incluyan tanto a técnicos en IA como a expertos en género, derecho y ciencias sociales.

Además, varios estudios enfatizan la importancia de la transparencia y auditoría periódica de las herramientas algorítmicas. Montesinos (2024) resaltaba que solo mediante un proceso de revisión constante es posible identificar y corregir el impacto que puede surgir en el uso de estas herramientas.

En cuanto a las recomendaciones para mitigar los efectos negativos, se sugiere que las herramientas no deben sustituir las decisiones humanas, sino que deben servir como apoyo complementario. En este sentido, Llorente (2022) y López (2024) subrayan que la automatización no debe desplazar el juicio humano en las decisiones judiciales, especialmente en las situaciones de alta vulnerabilidad. Faddoul *et al.* (2020) proponen el modelo Handoff Tree, una estrategia que incorpora un margen de error asociado a cada predicción algorítmica, lo que permite que la decisión final quede en manos del juez si la incertidumbre es demasiado alta, lo que podría reducir la disparidad predictiva entre razas y géneros. Finalmente, varios autores, como Barron (2007) y Montesinos (2023), proponen una formación crítica para los operadores del sistema judicial en el uso de estas herramientas, con el objetivo de evitar una lectura automática de los resultados y fomentar un análisis consciente de los posibles sesgos inherentes a los algoritmos.

3. Discusión

En las últimas décadas, la IA se ha comenzado a abrir camino en el sistema judicial y penitenciario, consolidándose como una herramienta que promete ayudar a mejorar los procesos y hacer que las decisiones sean más justas y objetivas, por ejemplo, en los sistemas de evaluación del riesgo (Van Dijck, 2022), en el análisis predictivo o en la toma de decisiones judiciales (Aletras *et al.*, 2016). Sin embargo, junto con estas promesas también aparecen dudas: ¿son realmente neutrales estos sistemas? Cada vez hay más evidencias de que, en lugar de eliminar desigualdades, pueden llegar a reforzarlas o incluso ampliarlas. Por ejemplo, algunos estudios han demostrado que el algoritmo COMPAS sobrestimaba el riesgo de reincidencia en personas negras y subestimaba el de personas blancas (Khademi y Honavar, 2020), que sistemas automatizados de decisión judicial pueden reproducir desigualdades socioeconómicas (Richardson *et al.*, 2019) y que algoritmos diseñados para ser objetivos pueden amplificar sesgos

preexistentes en los datos (Obermeyer *et al.*, 2019). Estos resultados resaltan la necesidad de evaluar críticamente los algoritmos judiciales e implementar mecanismos que permitan mitigar sesgos, para garantizar que la IA contribuya a decisiones más justas y equitativas.

Con ello, el análisis realizado a partir de la revisión sistemática pone de manifiesto que la incorporación de herramientas algorítmicas en el ámbito judicial y penitenciario es un fenómeno que, aunque responde a la búsqueda de una mayor objetividad y eficiencia, también plantea importantes retos sociales y éticos. En particular, se observa que la dimensión de género constituye un eje central para comprender las implicaciones de estas tecnologías en la toma de decisiones.

En este sentido, los artículos analizados tienden a señalar que las herramientas algorítmicas de valoración del riesgo podrían tener un impacto diferenciado según el género, reproduciendo e incluso amplificando desigualdades previamente existentes (López, 2024; Montesinos, 2023). Si bien gran parte de la literatura se ha centrado en los efectos sobre las mujeres cisgénero, una concepción más amplia sugiere que estos impactos podrían extenderse también a colectivos vulnerables, como personas trans, víctimas de violencia de género y jóvenes involucrados en procesos de justicia juvenil (Hamilton, 2022; López, 2024).

En general, los estudios revisados coinciden en señalar que estas herramientas tienden a reflejar o reforzar sesgos de género, en parte debido a la selección limitada o arbitraria de los datos de entrenamiento y a la falta de supervisión crítica en su diseño (Montesinos, 2023). Estas limitaciones podrían traducirse en decisiones desfavorables para las mujeres, particularmente en los casos de violencia de género o prisión preventiva (Hamilton, 2022; López, 2024). Algunos trabajos indican que estos sesgos no son meramente teóricos, sino que pueden materializarse en consecuencias concretas, como condenas más severas, mayor criminalización o menor acceso a medidas de protección, contribuyendo a mantener desigualdades estructurales (Barron, 2007).

El caso COMPAS, utilizado en el sistema judicial de Estados Unidos, se presenta como un ejemplo ilustrativo. Diversas investigaciones han sugerido que este sistema podría sobrestimar el riesgo en mujeres, derivando en penas más duras en comparación con hombres en circunstancias

similares (Blomberg *et al.*, 2010; Miró Llinars, 2018). En el litigio *Loomis vs. Wisconsin*, se identificaron posibles sesgos relacionados con el género y la raza, originados por un desequilibrio en los datos de entrenamiento, predominantemente masculinos, lo que habría resultado en una mayor precisión para los hombres y una tendencia a falsos positivos en el caso de las mujeres (Hamilton, 2022).

Este patrón parece vincularse con una representación desigual en los datos, que afectaría especialmente a mujeres jóvenes y racializadas. Flores (2022) plantea que los algoritmos tienden a patologizar conductas que, en varones, no se consideran problemáticas, reproduciendo estereotipos bajo una apariencia de objetividad técnica. Estas dinámicas podrían contribuir a procesos de exclusión social y a la perpetuación de la marginalidad de estos colectivos (Barron, 2007).

Las implicaciones trascienden el plano técnico y alcanzan dimensiones jurídicas y éticas, al incidir potencialmente en derechos fundamentales como el principio de igualdad o el derecho a una vida libre de violencia (Hamilton, 2022). En contextos de violencia de género, el uso de algoritmos sin una supervisión adecuada podría reforzar la estigmatización de las víctimas y consolidar visiones binarias o estereotipadas (Hamilton, 2022). Algunos autores advierten que esta desconfianza hacia los sistemas algorítmicos podría incluso desincentivar la presentación de denuncias (López, 2024).

Asimismo, la creciente automatización de la justicia ha sido señalada como un posible riesgo para la humanización del proceso judicial, tal como plantean autores como Basto (2024). Confiar en exceso en las recomendaciones algorítmicas podría limitar la capacidad del sistema para emitir decisiones contextualizadas y equitativas (Llorente, 2022; López, 2014). Si bien la tecnología puede contribuir a mejorar la eficiencia y sistematización de los casos, no debería sustituir el juicio humano, indispensable para abordar la complejidad y las particularidades de cada situación.

Un recurrente problema técnico identificado en la literatura es la confusión entre las categorías de sexo y género. Muchas herramientas incorporan la variable «sexo» (hombre o mujer) como criterio estadístico, sin considerar la dimensión social y relacional del género (Lamas, 2000; WHO, 2023). Esta simplificación binaria tiende a excluir a las personas no binarias o trans y omite factores estructurales como la victimización previa, la maternidad o la

desigualdad económica. El uso indiscriminado del sexo como variable fija puede generar distorsiones, como sobrestimaciones del riesgo en mujeres o evaluaciones inadecuadas para otros géneros invisibilizados (Barron, 2007). Si bien este enfoque busca objetividad, su aplicación sin una perspectiva de género resulta técnicamente limitada y éticamente cuestionable.

Por ello, y atendiendo a los resultados obtenidos a partir de la revisión de los artículos analizados, podemos concluir lo siguiente:

1. La adopción de algoritmos en el ámbito judicial y penitenciario busca mejorar la objetividad y la eficiencia en la toma de decisiones. No obstante, esto plantea importantes desafíos sociales y éticos, especialmente en relación con la equidad y la protección de los derechos fundamentales.
2. El género constituye una categoría clave para comprender el impacto de estas tecnologías. Los algoritmos tienden a reforzar sesgos de género debido a las limitaciones en los datos de entrenamiento y a la falta de supervisión crítica en su diseño, lo que afecta particularmente a mujeres cisgénero, personas trans, víctimas de violencia de género y jóvenes en procesos de justicia juvenil, entre otras.
3. La confusión entre sexo y género en las variables utilizadas por las herramientas algorítmicas puede generar evaluaciones inadecuadas o sobrestimaciones de riesgo, excluyendo a colectivos vulnerables y consolidando estereotipos.
4. La automatización creciente de la justicia puede limitar la capacidad de los jueces para emitir decisiones contextualizadas y equitativas. Si bien las herramientas algorítmicas contribuyen a la eficiencia y sistematización de los casos, no deben sustituir el juicio humano, indispensable para abordar la complejidad de cada situación.
5. En general, los sistemas algorítmicos no garantizan neutralidad absoluta. Tienden a reproducir y amplificar desigualdades estructurales de género, con el riesgo de consolidar estereotipos y vulnerar derechos fundamentales, incluyendo el principio de igualdad y el acceso a la justicia.

Sin embargo, es importante ser cautos con los resultados obtenidos en el presente estudio, atendiendo a una serie de limitaciones inherentes tanto al diseño metodológico adoptado como al estado actual de la producción científica en este ámbito. En primer lugar, la elección de una revisión sistemática como estrategia metodológica, si bien permite una síntesis rigurosa, transparente y reproducible de la evidencia disponible, implica que los hallazgos obtenidos dependen directamente de la cantidad, calidad y características de los estudios previamente publicados.

En este sentido, la principal limitación del estudio radica en la escasa disponibilidad de investigaciones empíricas que analicen de manera directa el impacto del género en el uso de herramientas algorítmicas de valoración del riesgo en los sistemas judicial y penitenciario. Aunque en la fase inicial de búsqueda se identificaron 54 publicaciones potencialmente relevantes, únicamente 7 cumplieron con los criterios de inclusión establecidos, lo que restringe el alcance analítico y la generalización de los resultados. No obstante, esta limitación no debe interpretarse como una debilidad metodológica del estudio, sino como un fiel reflejo del estado actual del arte en esta materia.

Asimismo, la evidencia disponible se caracteriza por una predominancia de estudios de carácter cualitativo y por una concentración geográfica en contextos occidentales, principalmente europeos y norteamericanos, lo que limita la posibilidad de extrapolar los hallazgos a otras jurisdicciones y marcos institucionales. Del mismo modo, la escasez de estudios con metodologías cuantitativas robustas o enfoques comparativos dificulta establecer con precisión el grado y la magnitud del impacto de género asociado al uso de estas herramientas algorítmicas en distintos contextos judiciales y penitenciarios. La reducida cantidad de estudios incluidos impide, además, realizar análisis más amplios sobre variaciones jurisdiccionales, tipologías de herramientas algorítmicas o diferencias según las distintas fases del proceso judicial o penitenciario. Para poder evaluar de manera más exhaustiva el impacto del género en estos contextos, habría sido necesario contar con una base empírica más amplia, circunstancia que actualmente no es posible debido a la limitada producción científica existente.

Estas limitaciones están estrechamente vinculadas al hecho de que la investigación sobre IA, justicia y género se encuentra aún en una fase incipiente, en la que pre-

domina un debate teórico y normativo sobre los sesgos algorítmicos y la discriminación automatizada, frente a estudios empíricos sistemáticos que evalúen sus efectos concretos en la toma de decisiones judiciales. En este sentido, la escasez de literatura empírica constituye no solo una limitación, sino también un hallazgo relevante de la propia revisión, al evidenciar importantes vacíos en el conocimiento disponible.

Pese a estas restricciones, los siete estudios incluidos permiten realizar un primer mapeo sistemático de la evidencia científica existente y ofrecen una aproximación fundamentada a las principales líneas de impacto de género identificadas hasta el momento. No obstante, la limitada cantidad de investigaciones disponibles aconseja interpretar los resultados con cautela y evita extraer conclusiones generalizables más allá del alcance exploratorio del estudio. Es por ello por lo que este trabajo debe entenderse como una exploración inicial del campo, orientada a sistematizar la evidencia disponible y a identificar tendencias, riesgos y debates emergentes, más que como una evaluación exhaustiva del impacto del género en el uso de herramientas algorítmicas. A partir de estos resultados, futuras líneas de investigación deberían orientarse hacia el desarrollo de estudios empíricos más sólidos, que incorporen metodologías cuantitativas, diseños comparativos entre jurisdicciones, análisis longitudinales y evaluaciones específicas de herramientas concretas de valoración del riesgo.

Finalmente, resulta especialmente relevante promover investigaciones interdisciplinarias que integren perspectivas técnicas, jurídicas y sociales, con la finalidad de evaluar no solo la existencia de sesgos de género, sino también la eficacia real de las estrategias de mitigación propuestas. En este sentido, el presente estudio pone de manifiesto la necesidad de que futuras investigaciones revisen y complementen estos hallazgos mediante metodologías más robustas, tal y como reconoce la propia literatura analizada.

Reconocimientos

La publicación es parte del proyecto (lus_machina) TED2021-129356B-I00, financiado por MCIN/AEI/10.13039/501100011033 y por la Unión Europea Next-GenerationEU/PRTR.

Referencias bibliográficas

- ALETRAS, N.; TSARAPATSANIS, D.; PREOȚIUC-PIETRO, D.; LAMPOS, V. (2016). «Predicting judicial decisions of the European Court of Human Rights: A natural language processing perspective». *PeerJ Computer Science*, vol. 2, e93. DOI: <https://doi.org/10.7717/peerj-cs.93>
- AN, J.; HUANG, D.; LIN, C.; TAI, M. (2025). «Measuring gender and racial biases in large language models: Intersectional evidence from automated resume evaluation». *PNAS Nexus*, vol. 4, n.º 3. DOI: <https://doi.org/10.1093/pnasnexus/pgaf089>
- BARONA, S. (2019). «Inteligencia artificial o la algoritmización de la vida y de la justicia: ¿solución o problema?», *Revista Boliviana de Derecho*, n.º 28, Universidad Autónoma Gabriel René Moreno, Santa Cruz de la Sierra. Bolivia. pp. 18-49.
- BARRON, C. (2007). *Governing girls: Rehabilitation in the age of risk*. Tesis doctoral. Simon Fraser University.
- BASTO, B. (2024). «Impacto de la inteligencia artificial en la argumentación jurídica: Un nuevo paradigma de la administración de justicia». *Argumentación jurídica e inteligencia artificial*.
- BLOMBERG, T.; BALES, W.; MANN, K.; MELDRUM, Y. NEDELEC, J. (2010). *Validation of the COMPAS risk assessment Classification instrument 4*. Center for Criminology and Public Policy Research. College of Criminology and Criminal Justice.
- BUOLAMWINI, J.; GEBRU, T. (2018). «Gender Shades: Intersectional Accuracy Disparities in Commercial Gender Classification». *Proceedings of Machine Learning Research*, vol. 81, págs. 77-91 [en línea]. Disponible en: <https://proceedings.mlr.press/v81/buolamwini18a/buolamwini18a.pdf>
- BUSTELO, J. L. (2025). «Hacia una gobernanza algorítmica transparente: auditoría de sesgo. Estudio de caso». *Cuadernos de Gobierno y Administración Pública*, vol. 12, n.º 1, págs. 27-39. Ediciones Complutense. DOI: <https://doi.org/10.5209/cgap.97604>
- CASANOVAS, P.; PALAO, J. M. (2019). «Inteligencia Artificial y Derecho: Retos y oportunidades». *Revista de Derecho Informático*, vol. 23, págs. 1-24.
- CASANOVAS, P.; RODRÍGUEZ-DONCEL, V.; GONZÁLEZ-CONJERO, J. (2017). «The Role of Pragmatics in the Web of Data». En: Poggi, F.; Capone, A. (eds.). *Pragmatics and Law. Perspectives in Pragmatics, Philosophy & Psychology*, vol. 10. Cham: Springer. DOI: https://doi.org/10.1007/978-3-319-44601-1_12
- CENTENO-ROMERO M.; CHACÓN-ARRIETA G.; VEGA-AGUILARJA; GONZALEZ-TORRES, A., LEITÓN-JIMÉNEZ, J. (2021). «Revisión sistemática de literatura: aplicaciones de las comunicaciones moleculares». *Revista Tecnología En Marcha*, vol. 34, n.º 2, págs. 147-160. DOI: <https://doi.org/10.18845/tm.v34i2.4982>
- CENTRO COCHRANE IBEROAMERICANO (TRADUCTORES). (2012). *Manual Cochrane de Revisiones Sistemáticas de Intervenciones* (Versión 5.1.0). Barcelona: Centro Cochrane Iberoamericano [en línea]. Disponible en: <http://www.cochrane.es/?q=es/node/269>.
- COMISIÓN EUROPEA. (2020). *Libro Blanco sobre la inteligencia artificial: Un enfoque europeo orientado a la excelencia y la confianza*. Bruselas: Comisión Europea.
- COOPER, H.; HEDGES, L. (1994). *The Handbook of Research Synthesis*. Nueva York: Russell Sage Foundation.
- CORVALAN, J.; MACCHIAVELLI, N.; CINALLI, C.; DANESI, C.; DÁVILA, L.; TOLOSA, P.; SÁNCHEZ CAPARRÓS, M.; MAFUD, M. (2021). «Sesgos algorítmicos de género». *IALAB/Universidad de Buenos Aires* [en línea]. Disponible en: https://www.researchgate.net/publication/359115499_Sesgos_algoritmicos_de_genero
- CRIADO PÉREZ, C. (2019). *Invisible Women: Data Bias in a World Designed for Men*. Londres: Chatto & Windus.
- FADDIOL, M.; RUHRMANN, H.; LEE, J. (2020). «A risk assessment of a pretrial risk assessment tool: Tussles, mitigation strategies, and inherent limits». *arXiv*. <https://doi.org/10.48550/arXiv.2005.07299>

- FARFÁN INTRIAGO, J. L.; FARFÁN LARGACHA, J. A.; FARFÁN LARGACHA, B.; NÚÑEZ VERA, J. P. (2023). «Inteligencia artificial y Derecho: ¿La justicia en manos de la IA?». *FRÓNESIS: Revista de Filosofía Jurídica, Social y Política*, vol. 30, n.º 2, págs. 173-197.
- FERREIRA, I.; URRÚTIA, G.; ALONSO-COELLO, P. (2011). «Revisiones sistemáticas y metaanálisis: bases conceptuales e interpretación». *Revista Española de Cardiología*, vol. 64, n.º 8, págs. 688-696. DOI: <https://doi.org/10.1016/j.recesp.2011.03.029>
- FLORES, J. (2022). *Ética y prospectiva de la inteligencia artificial en la educación*. Universidad de Salamanca.
- HAMILTON, M. (2022). «Modelling Pretrial Detention». *American University Law Review*, vol. 7. DOI: <https://doi.org/10.2139/ssrn.4188907>
- HARCOURT, B. E. (2015). «Risk as a Proxy for Race. The Dangers of Risk Assessment». *Federal Sentencing Reporter*, vol. 27, n.º 4, págs. 237-243. DOI: <https://doi.org/10.1525/fsr.2015.27.4.237>
- KHADEMI, A.; HONAVAR, V. (2020). «Algorithmic Bias in Recidivism Prediction: A Causal Perspective (Student Abstract)». *Proceedings of the AAAI Conference on Artificial Intelligence*, vol. 34, n.º 10, págs. 13839-13840. DOI: <https://doi.org/10.1609/aaai.v34i10.7192>
- LAMAS, M. (2000). «Diferencias de sexo, género y diferencia sexual». *Cuicuilco Revista de Ciencias Antropológicas*, vol. 7, n.º 18, págs. 95-118. Disponible en: <https://revistas.inah.gob.mx/index.php/cuicuilco/article/view/360>
- LANSING, S. (2012). «New York State COMPAS-Probation Risk and Need Assessment Study: Examining the Recidivism Scale's Effectiveness and Predictive Accuracy». *U.S. Department of Justice. Office of Justice Programs* [en línea].
- LARSON, J.; MATTU, L. K.; AMGWIN, J. (2016). «How We Analyzed the COMPAS Recidivism Algorithm». *ProPublica* [en línea]. <https://www.propublica.org/article/how-we-analyzed-the-compas-recidivism-algorithm>
- LLORENTE, M. (2022). «Hacia una justicia penal predictiva». *Cuadernos de Política Criminal*, n.º 136, págs. 91-124.
- LONG, H.; FRENCH, D.; BROOKS, J. (2020). «Optimising the value of the critical appraisal skills programme (CASP) too for quality appraisal in qualitative evidence synthesis». *Sage Journals*. DOI: <https://doi.org/10.1177/2632084320947559>
- LÓPEZ DE ZUBIRÍA DÍAZ, S. (2024). «El uso de herramientas de inteligencia artificial en materia de violencia de género: ¿una adecuada estrategia preventiva y represiva?». *Revista Penal México*, vol. 13, n.º 25, págs. 213-236. DOI: <https://doi.org/10.29166/cap.v7i13.5931>
- MAYSON, S. G. (2018). «Bias in, bias out». *Yale Law Journal*, vol. 128, págs. 2218-2300.
- MIRÓ LLINARES, F. (2018). «Inteligencia artificial y justicia penal: más allá de los resultados lesivos causados por robots». *Revista de Derecho Penal y Criminología*, 3.ª Época, n.º 20, págs. 87-130. UNED. DOI: <https://doi.org/10.5944/rdpc.20.2018.26446>
- MONTESINOS, A. (2023). «Algoritmos predictivos y perspectiva de género en el proceso penal». En: Miró, F. (coord.). «Digitalización y algoritmización de la justicia», págs. 39-59. *IDP. Revista de Internet, Derecho y Política*, n.º 39. DOI: <https://doi.org/10.7238/idp.v0i39.416439>
- MONTESINOS, A. (2024). «Inteligencia artificial en la justicia con perspectiva de género: amenazas y oportunidades». *Actualidad Jurídica Iberoamericana*, vol. 21, págs. 566-597.
- MOHER, D.; LIBERATI, A.; TETZLAFF, J.; ALTMAN, D. G. (2009). «Preferred reporting items for systematic reviews and meta-analyses: The PRISMA statement». *PLOS Medicine*, vol. 6, n.º 7, e1000097. DOI: <https://doi.org/10.1371/journal.pmed.1000097>
- OBERMEYER, Z.; POWERS, B.; VOGELI, C.; MULLAINATHAN, S. (2019). «Dissecting racial bias in an algorithm used to manage the health of populations». *Science*, vol. 366, n.º 6464, págs. 447-453. DOI: <https://doi.org/10.1126/science.aax2342>

- PÁGE, M. J.; MCKENZIE, J. E.; BOSSUYT, P. M.; BOUTRON, I.; HOFFMANN, T. C.; MULROW, C. D.; SHAMSEER, L.; TETZLAFF, J. M.; AKL, E. A.; BRENNAN, S. E. *et al.* (2021). «Declaración PRISMA 2020: una guía actualizada para la publicación de revisiones sistemáticas». *Revista Española de Cardiología*, vol. 74, págs. 790-799. DOI: <https://doi.org/10.1016/j.rec.2021.07.010>
- PARDAL-REFOYO, J. L.; PARDAL-PELÁEZ, B. (2020). «Anotaciones para estructurar una revisión sistemática». *Revista ORL*, vol. 11, n.º 2, págs. 155-160. DOI: <https://doi.org/10.14201/orl.22882>
- PASQUALE, F. (2015). *The Black Box Society: The Secret Algorithms That Control Money and Information*. Cambridge: Harvard University Press. DOI: <https://doi.org/10.4159/harvard.9780674736061>
- PÉREZ DOMÍNGUEZ, S. (2024). «Análisis de factores situacionales relacionados con la violencia de género: una revisión sistemática desde la criminología ambiental». *Revista General de Derecho Penal*, vol. 42.
- PÉREZ DOMÍNGUEZ, S.; CASTRO-TOLEDO, F. J.; MIRÓ-LLINARES, F. (2019). «Prevalencia, factores asociados y diferencias de género en el cumplimiento de la propiedad intelectual: una revisión sistemática». *Revista Electrónica de Criminología*, 02-04.
- RICHARDSON, R.; SCHULTZ, J.; CRAWFORD, K. (2019). «Dirty data, bad predictions: How civil rights violations impact police data, predictive policing systems, and justice». *NYU Law Review Online*, vol. 94, págs. 192-233.
- REINA, M. (2019). «Análisis del fenómeno de la trata de personas desde la perspectiva de los derechos humanos». *Revista Cuatrimestral de las Facultades de Derecho y Ciencias Económicas y Empresariales*.
- SAAVEDRA VERA, C. O.; JÁUREGUI BUSTAMANTE, K. DEL R.; ARISTA BUSTAMANTE, L. L. (2023). «La incidencia del sesgo algorítmico en la justicia predictiva del sistema judicial». *TZHOECOEN*, vol. 15, n.º 2, págs. 79-97. DOI: <https://doi.org/10.26495/tzh.v15i2.2592>
- SALCIDO, M.; VARGAS, A.; MEDINA, N.; RAMÍREZ, F.; GARCÍA, M.; BRISEÑO, A.; JIMÉNEZ, J. (2021). «Revisión sistemática: el más alto nivel de evidencia». *Orthotips*, vol. 17, n.º 4, págs. 217-221. DOI: <https://dx.doi.org/10.35366/102220>
- SARRIDIS, A.; SMITH, J.; KIM, L. (2023). «Towards Fair Face Verification: An In-depth Analysis of Demographic Biases». *arXiv* [en línea]. Disponible en: <https://arxiv.org/abs/2307.10011>
- SCHONNING, P. (1996). «Applicable Law in Transfrontier on B-Line Transmissions». *Revue Internationale du Droit d'Auteur*, vol. 170, págs. 21 y ss.
- STARR, S. B. (2014). «Evidence-based sentencing and the scientific rationalization of discrimination» [en línea]. *Stanford Law Review*, vol. 66, págs. 803-872. Disponible en: http://www.stanfordlawreview.org/wp-content/uploads/sites/3/2014/04/66_Stan_L_Rev_803-Starr.pdf
- UNESCO (2020). «Artificial intelligence and gender equality: Key findings of UNESCO's Global Dialogue». UNESCO. Disponible en: <https://unesdoc.unesco.org/ark:/48223/pf0000374174>
- VAN DIJCK, G. (2022). «Predicting Recidivism Risk Meets AI Act». *European Journal on Criminal Policy and Research*, vol. 28, n.º 3, págs. 407-423. DOI: <https://doi.org/10.1007/s10610-022-09516-8>
- WACHTER-BOETTCHER, S. (2017). *Technically Wrong: Sexist Apps, Biased Algorithms, and Other Threats of Toxic Tech*. Nueva York: W.W. Norton & Company.
- WORLD HEALTH ORGANIZATION (WHO). (2023). *Gender and Health*. Ginebra: WHO.
- XENIDIS, R.; SENDEN, L. (2020). «EU non-discrimination law in the era of artificial intelligence: Mapping the challenges of algorithmic discrimination». En: Bernitz, U., Madsen, M. R., Petersen, N.; Weatherill, S. (eds.). *General Principles of EU Law and the EU Digital Order*, págs. 151-182. Kluwer Law International.

Cita recomendada

MARTÍNEZ ALMANZA, Rocío; PÉREZ DOMÍNGUEZ, Sandra (2026). «Género y herramientas algorítmicas: una revisión sistemática sobre los retos y el impacto que plantean en el ámbito judicial y penitenciario». *IDP. Revista de Internet, Derecho y Política*, núm. 45. UOC [Fecha de consulta: dd/mm/aa]. DOI: <http://dx.doi.org/10.7238/idp.v0i45.9800659>



Los textos publicados en esta revista están –si no se indica lo contrario– bajo una licencia Reconocimiento-Sin obras derivadas 3.0 España de Creative Commons. Puede copiarlos, distribuirlos y comunicarlos públicamente siempre que cite su autor y la revista y la institución que los publica (*IDP. Revista de Internet, Derecho y Política*; UOC); no haga con ellos obras derivadas. La licencia completa se puede consultar en: <http://creativecommons.org/licenses/by-nd/3.0/es/deed.es>.

Sobre las autoras**Rocío Martínez Almanza**

Centro CRÍMINA, Universidad Miguel Hernández de Elche
r.martinezalmanza@crimina.es

Profesora ayudante doctora en la Universidad Miguel Hernández de Elche (UMH), en el área de Ciencias Políticas y de la Administración, e investigadora en el Centro de Investigación CRÍMINA para el estudio y prevención de la delincuencia. Es doctora en Criminología por la UMH, donde realizó su tesis sobre *Actitudes y percepciones hacia la criminalización de la prostitución*, cuenta con un máster en Intervención Criminológica y Victimológica y en Evaluación de Políticas Públicas. Ha sido investigadora predoctoral en el Centro CRÍMINA, ha realizado estancias de investigación en la Universidad Federal de Minas Gerais (2023) y en la Universidad de Oporto (2024). Ha participado en y forma parte de proyectos competitivos, como FAKEPENAL, lusMachina, GamerVictim, Tesserat o Success, entre otros. Es autora de artículos sobre sus líneas principales de investigación. Ha impartido docencia en diferentes grados y actualmente es docente y dirige trabajos de fin de grado y de máster en el Grado en Seguridad Pública y Privada (UMH), Grado en Ciencias Políticas y Gestión Pública (UMH), Máster en Cibercriminalidad (UOC), Máster en Análisis de Prevención del Crimen (UMH) y Máster en Gerontología y Salud (UMH).

Sandra Pérez Domínguez

Centro CRÍMINA, Universidad Miguel Hernández de Elche
s.perezd@crimina.es

Profesora en el Departamento de Psicología de la Salud de la Universidad Miguel Hernández de Elche, en el área de Psicología Evolutiva y de la Educación. Es graduada en Psicología por la misma universidad y cuenta con un máster en Intervención Criminológica y Victimológica. Asimismo, es doctora por la Universidad Miguel Hernández, cuya tesis se centró en el papel del factor humano en el proceso de transformación digital 4.0 del sistema de justicia penal, y que cuenta con mención internacional e industrial. Forma parte del Grupo de Investigación Tecnología, Mente y Comportamiento Social y Desviado y del centro Crímina, donde ha participado en diferentes proyectos nacionales e internacionales como lusMachina, GamerVictim, Tesserat o Success. Imparte docencia en diferentes grados y actualmente es docente y dirige trabajos de fin de grado y de máster en el Grado en Psicología (UMH), el grado en Criminología (UOC), el máster en Cibercriminalidad (UOC), el máster en Análisis de Prevención del Crimen (UMH), el máster en Gerontología y Salud (UMH) y el máster en Intervención Criminológica y Victimológica (UMH).