

Inteligencia artificial en la seguridad pública y en el sistema penal en América Latina

Análisis basado en el debido proceso



Con el respaldo de **moz://a**

Sobre Fair Trials

Fair Trials, Juicio Justo en español, es una organización internacional de la sociedad civil dedicada a monitorear los sistemas penales alrededor del mundo. La organización trabaja por la consecución de sistemas de justicia criminal justos e igualitarios, con un equipo compuesto por expertos de diferentes regiones del mundo. Fair Trials ha seguido y analizado el crecimiento del uso de inteligencia artificial en el marco del sistema penal desde 2020. En este informe trabajó Luis Eliud Tapia, Consultor Legal, bajo la coordinación de Verónica Hinestroza, Asesora Jurídica Senior.

Contacto

Verónica Hinestroza

Asesora Jurídica Senior

Veronica.Hinestroza@fairtrials.net

 @fairtrials  @fairtrials  Fair Trials

First published by Fair Trials in December 2024. © Fair Trials 2024.

“Fair Trials” includes Fair Trials International, Fair Trials Europe, and Fair Trials Americas. Fair Trials International is a registered charity (no. 1134586) and in 2010 was incorporated with limited liability in England and Wales (No. 7135273), and is based at 5 Castle Road, London, NW1 8PR. In May 2014, Fair Trials International founded Fair Trials Europe, which is a registered public foundation in Belgium (registered number 0552.688.677). In 2018 we founded Fair Trials Americas, which is a registered 501(c)(3) public charity in the United States of America (No DLN17053243307017). We were initially founded in 1992 with the name “Fair Trials Abroad”.

Agradecimientos

Fair Trials agradece el apoyo de la Fundación Mozilla para la realización del estudio en que se basa este informe.

Extendemos nuestro agradecimiento a las organizaciones socias involucradas en la organización del Congreso Internacional Inteligencia Artificial, Derecho y Democracia, que tuvo lugar en Ciudad de México en octubre de 2023: la Universidad Iberoamericana Ciudad de México, la Fundación Konrad Adenauer y el Centro de Estudios Constitucionales de la Suprema Corte de Justicia de la Nación en México. Nuestra gratitud a los y las ponentes que participaron en el encuentro y que con sus reflexiones aportaron a las conclusiones de este informe, y a Claudia Pajas Tinoco, quién desde su investigación como pasante de Fair Trials, apoyó la realización del estudio.

1. Introducción	5
2. Resumen ejecutivo	6
Introducción	6
Europa	7
América Latina	7
Desafíos y riesgos	8
Conclusiones	9
Recomendaciones	9
3. Conceptos básicos	10
4. Inteligencia artificial en labores de seguridad pública y por los sistemas de justicia penal en Europa	11
5. Conclusiones preliminares	14
6. Inteligencia artificial en labores de seguridad pública y en la administración de la justicia penal en América Latina	15
7. Inteligencia artificial y desafíos emergentes en América Latina frente al debido proceso	23
8. Conclusiones	27
9. Recomendaciones	28

1. Introducción

El uso de los sistemas basados en inteligencia artificial (IA), en la seguridad pública y en el sistema penal en América Latina responde a una tendencia global, pero encuentra además explicación en dinámicas propias de la región. En general, en América Latina, los niveles de violencia e impunidad llevan a que la población ejerza presión sobre las autoridades. Los gobiernos tienden a responder a esta presión con medidas que aumenten tanto la protección directa frente al crimen como la percepción de seguridad. Sin embargo, buscar integrar nuevas tecnologías para responder a preocupaciones legítimas, no exime a los Estados de garantizar los derechos humanos en general, y en especial aquellos relacionados con el debido proceso.

Organizaciones de la sociedad civil alrededor del mundo han documentado los impactos en los derechos humanos de la tecnología basada en la inteligencia artificial, incluyendo, por ejemplo, el reconocimiento facial. Los derechos primeramente vulnerables a dicha tecnología son los derechos a la privacidad y a la igualdad y no discriminación.

Sin embargo, al estudiar el sistema de justicia en su conjunto, es posible constatar que la inteligencia artificial impacta otros derechos humanos como la libertad personal, la presunción de inocencia y el debido proceso.

Este informe analiza el uso de sistemas basados en inteligencia artificial en el sistema de justicia bajo el paraguas del debido proceso; un campo de estudio en desarrollo. El informe toma como precedente algunas experiencias documentadas por Fair Trials en países de Europa para centrarse en adelante en América Latina. Las fuentes analizadas para el análisis regional incluyen insumos de la academia, sociedad civil, organismos internacionales y cortes de justicia.

Adicionalmente, el informe se nutre de las discusiones que tuvieron lugar entre el 23 y 24 de octubre de 2023, en el Congreso Internacional Inteligencia Artificial, Derecho y Democracia, organizado en Ciudad de México por Fair Trials, la Universidad Iberoamericana Ciudad de México, la Fundación Konrad Adenauer y el Centro de Estudios Constitucionales de la Suprema Corte de Justicia de la Nación en México. Durante 2 días, se conversó en 5 paneles sobre oportunidades y retos de la inteligencia artificial, con un enfoque en seguridad, justicia, discriminación, desigualdad y democracia.

En la mesa sobre seguridad e inteligencia artificial participaron Verónica Hinestroza de Fair Trials, Marc Rotenberg del Center for Artificial Intelligence and Digital Police (CAIDP), Priscilla Ruiz de Article 19 y Pablo Nunes de Panóptico. En la mesa sobre justicia e inteligencia artificial participaron Carolina Villadiego de la Comisión Internacional de Juristas (CIJ), Grecia Macías de la Red en Defensa de los Derechos Digitales (R3D), Carlos Guzmán, Magistrado en materia Penal del Tribunal Superior de Bogotá y Luis Tapia de Fair Trials.

El informe tiene como eje central reafirmar que las salvaguardas del juicio justo o debido proceso no son renunciables, especialmente frente al uso de tecnologías basadas en inteligencia artificial. Tampoco el uso eficiente de recursos justifica relativizar las garantías mínimas del debido proceso, cuyo origen buscó limitar el ejercicio arbitrario del poder público.

Cabe destacar, sin embargo, que los poderes judiciales, las fiscalías y las instituciones de seguridad están utilizando IA para otros fines como la gestión judicial, desarrollo de

buscadores de jurisprudencia y para organizar la información que recopilan¹.

En este sentido, el informe reconoce que resulta fundamental identificar y diferenciar los posibles usos de la inteligencia artificial en el sistema de justicia para evitar descalificar la tecnología a priori. El informe toma en cuenta que la IA puede ser útil para el funcionamiento de los poderes judiciales, y genera, a su vez, una alerta ante el llamado solucionismo tecnológico², destacando que resulta imperante desarrollar mecanismos que regulen su uso con un enfoque de derechos humanos, evitando adoptar esta tecnología acríticamente sacrificando los derechos de las personas.

Por ello, el informe busca identificar tanto riesgos como buenas prácticas relacionadas con el uso de la inteligencia artificial por las autoridades encargadas de la seguridad pública, de los sistemas de justicia penal y de los poderes judiciales en América Latina y Europa. Además, pretende reflexionar sobre la reconfiguración de las garantías del debido proceso en clave de inteligencia artificial.

Se espera que el informe sea un insumo que nutra una conversación necesaria para la garantía de la dignidad humana desde la seguridad y los derechos humanos dado el uso creciente de la inteligencia artificial por las autoridades del policíacas y del sistema de justicia.

2. Resumen ejecutivo

Introducción

El informe analiza el uso de la inteligencia artificial (IA) en la seguridad pública y en el sistema penal en América Latina, destacando la necesidad de garantizar los derechos humanos, especialmente el debido proceso. El informe parte de experiencias documentadas en Europa, y recupera, entre otros debates, discusiones del Congreso Internacional "Inteligencia Artificial, Derecho y Democracia" realizado en la Universidad Iberoamericana el 23 y 24 de octubre de 2023.

A partir de lo anterior, se exponen experiencias de Argentina, Brasil, Chile, Colombia, El Salvador, México y Uruguay respecto al uso de IA en la seguridad pública, en el sistema penal y por los poderes judiciales.

Hallazgos principales en materia de derechos humanos de la experiencia del uso de algunos sistemas de IA

Europa

1) Seguridad pública

- a) **RADAR-iTE (Alemania)**. Evaluación de riesgo de violencia con un enfoque en potenciales terroristas islamistas, evidenciando sesgos contra musulmanes.
- b) **ProKid (Holanda)**. Predicción de comisión de delitos por parte de niños, niñas y adolescentes, con un alto índice de errores y sesgos.

¹ Carolina Villadiego. Mesa sobre justicia e inteligencia artificial del Congreso Internacional Inteligencia Artificial, Derecho y Democracia organizado por la Universidad Iberoamericana Ciudad de México y la Fundación Konrad Adenauer, Fair Trials octubre 2023.

² Luis Fernando García y Grecia Macías, La nueva cara de la discriminación: el impacto diferenciado de la tecnología de reconocimiento facial en grupos vulnerables en SCJN, Discriminación, Problemas Contemporáneos, Ricardo Latapie Aldana (Coord.), págs. 109 y 110.

- c) **NDAS (Reino Unido)**. Evaluación de riesgos que mostró discriminación racial.
- d) **HART (Reino Unido)**. Predicción de reincidencia con una precisión limitada y uso de datos discriminatorios.

2) Justicia penal y sistemas penitenciarios

- a) **Cassandra (Ucrania)**. Evaluación de riesgo de reincidencia, con falta de transparencia en los datos impactando el derecho a acceder a la información y al debido proceso.
- b) **RisCanvi (España)**. Evaluación de libertad anticipada con dificultades para apelar decisiones influidas por el diseño del sistema.

América Latina

1) Seguridad pública

- a) **Reconocimiento Facial**. Uso extensivo con problemas de sesgos y falsos positivos que afectan a poblaciones en situación de vulnerabilidad.
 - i) **Ol. (Brasil)**. La tecnología de reconocimiento facial facilita a la policía justificar su actuación alegadamente ilegal y tiene un sesgo discriminatorio racial hacia personas afrodescendientes debido a la zona donde se instalan las cámaras.
 - ii) **Policía de Proximidad. (México)**. Mediante cámaras de vigilancia se discrimina a personas de bajos recursos y de la comunidad LGTBTTQ+. Incorporar reconocimiento facial a cámaras de vigilancia puede reproducir los sesgos de las personas y facilitar la discriminación.
 - iii) **Calle Segura. (Chile)**. El uso de reconocimiento facial en drones es opaco, altamente falible e institucionaliza sesgos asociados a la clase social y el color de piel.
 - iv) **Reconocimiento facial. (El Salvador)**. El uso de reconocimiento facial por gobiernos acusados de cometer graves abusos a los derechos humanos puede ser altamente contraproducente contra poblaciones estigmatizadas y discriminadas.
- b) **Predicción de delitos**. Implementación limitada y con precauciones que muestra sesgos y errores en las predicciones.
 - i) **Sistema Predictivo del Delito Urbano. (Chile)**. Sistema que buscaba predecir en qué zonas se cometerán delitos basado en algoritmos. El sistema puede facilitar que la policía hostigue a personas por considerarlas sospechosas.
 - ii) **Estudio empírico de identificación de comportamiento criminal. (Chile)**. Investigación que se centró en la posibilidad de predecir el riesgo de reincidencia usando algoritmos. En más del 30% de los casos las predicciones fueron erradas. Además, se identificó que el sistema puede alentar la detención de personas por sus antecedentes y no por conductas cometidas.
 - iii) **Predicción del delito en ciudades intermedias. (Colombia)**. Los modelos de predicción de delitos pueden tener efectos discriminatorios contra poblaciones o personas que viven en zonas específicas. Además, dichos sistemas cometen errores que se profundizan cuando hay poca información como en ciudades intermedias como Bucaramanga.

- iv) PredPol. (Uruguay). Herramienta que buscaba predecir dónde se cometerían delitos. Un programa piloto encontró que el sistema no mostraba mejores resultados que las tecnologías tradicionales.
- 2) Poderes judiciales y justicia Penal. En el sistema penal el uso de la inteligencia artificial está en exploración o se ha descartado por sus riesgos. En las labores generales del sistema de justicia, se están implementando algunos sistemas basados en IA para apoyar la elaboración de sentencias.
- a) Valoración de pruebas y sentencias. (México). Ha sido explorada la posibilidad de utilizar inteligencia artificial en el dictado de sentencias en la revista del Instituto Nacional de Ciencias Penales pero hasta la fecha no se ha implementado.
 - b) PRiSMA (Colombia). Desarrollo un sistema de evaluación de riesgo de reincidencia basado en aprendizaje automatizado con preocupaciones sobre la afectación al derecho a la igualdad y no discriminación, al debido proceso, al derecho de defensa y a la imparcialidad judicial.
 - c) Projeto Sócrates (Brasil). Sistema basado en inteligencia artificial que busca reducir los tiempos para emitir sentencias analizando sentencias previas que sean similares. Sin embargo, se corre el riesgo de automatizar injusticias históricas del sistema judicial.
 - d) PROMETEA (Argentina). Sistema basado en inteligencia artificial para automatizar la elaboración de dictámenes jurídicos. Implicó una reducción dramática del tiempo de resolución de algunos casos. 90 minutos a 1 minuto en los pliegos de contrataciones. El sistema incorpora controles basados en derechos humanos.
 - e) PretorIA (Colombia). Herramienta basada en inteligencia artificial que ayuda a preseleccionar las tutelas que serán revisadas. En Colombia se reciben alrededor de 600,000 tutelas por año. Si bien es un sistema que automatiza actividades la decisión es tomada por una persona a fin de reducir riesgos para los derechos humanos.

Desafíos y riesgos

- Libertad Personal: Riesgo de restricciones arbitrarias a la libertad deambulatoria de personas respecto de las cuales no se ha comprobado que han cometido ningún delito.
- Debido Proceso: Afectación a la presunción de inocencia, derecho de defensa, derecho a recurrir el fallo, y la independencia e imparcialidad judicial debido a que, entre otras cosas, no se conoce la forma en que se toma la decisión, de qué bases de datos proviene la información quién financia los sistemas basados en inteligencia artificial.
- Igualdad y no discriminación. La tecnología basada en inteligencia artificial automatiza los sesgos discriminatorios que caracterizan las labores de seguridad pública y del sistema de justicia penal.

Conclusiones

El uso de IA en la justicia y seguridad en Europa ha mostrado un alto costo para los derechos humanos, con sesgos y errores significativos. En América Latina, el uso es más limitado pero presenta riesgos similares. Es necesario un enfoque equilibrado que no condene ni adopte acríticamente la IA, asegurando siempre la protección de los derechos humanos, especialmente el debido proceso.

Se subraya la importancia de un enfoque regulado y consciente de los riesgos que la inteligencia artificial puede presentar en el ámbito de la seguridad pública y la justicia,

promoviendo siempre la protección de los derechos humanos de las personas, especialmente, el derecho al debido proceso.

Existen algunas oportunidades favorables para el uso de IA en los sistemas de justicia como ejemplifican los casos de Argentina y Colombia siempre que se privilegie la transparencia, la rendición de cuentas y el respeto a los derechos humanos.

Recomendaciones

- 1) Prohibiciones:
 - a) Uso de tecnologías de vigilancia masiva como el reconocimiento facial.
 - b) Uso de sistemas de predicción de delitos.
- 2) Transparencia y rendición de cuentas:
 - a) Informar sobre la adquisición y uso de sistemas de IA.
 - b) Desarrollar garantías del debido proceso aplicables a sistemas de IA.
- 3) Buenas prácticas:
 - a) Realizar pruebas piloto para evaluar el funcionamiento y riesgos antes de implementar los sistemas.
 - b) Implementar auditorías externas a las tecnologías de IA.
- 4) Oportunidades controladas:
 - a) Uso de sistemas de IA con salvaguardas adecuadas, siguiendo ejemplos de Argentina y Colombia.

3. Conceptos básicos

- **Inteligencia artificial**

De acuerdo con la Organización de las Naciones Unidas (ONU), la Inteligencia Artificial (IA) se define como una "constelación de procesos y tecnologías que permiten que las computadoras complementen o reemplacen tareas específicas que de otro modo serían ejecutadas por seres humanos, como tomar decisiones y resolver problemas"¹.

El académico de la Universidad de Buenos Aires, Juan Corvalán, en el Tratado de Inteligencia Artificial y Derecho refiere 20 definiciones sobre inteligencia artificial². Ello refleja que el concepto de IA está en constante debate y en construcción³.

La organización Access Now refiere que la IA es considerada más un campo de estudio que una cosa y que puede desdoblarse en subcampos de estudios como el aprendizaje automatizado y la robótica⁴.

Una definición adicional de IA refiere que es "un conjunto de procesos y técnicas que

¹ Res. 73/348 de la Asamblea General de la ONU citado en Juan Corvalán et. al. Inteligencia artificial: bases conceptuales para una aproximación interdisciplinar en Juan Corvalán, Tratado de Inteligencia Artificial y Derecho, T. 1.

² Juan Corvalán et. al. Inteligencia artificial: bases conceptuales para una aproximación interdisciplinar en Juan Corvalán, Tratado de Inteligencia Artificial y Derecho, T. 1.

³ Access Now. Human Rights in the Age of Artificial Intelligence, pág 8.

⁴ Id.

permiten a un sistema computacional hacer cálculos estadísticos capaces de identificar patrones y a partir de ellos, predecir comportamientos¹". Sin embargo, la IA "está lejos de tener una capacidad crítica o empática²".

La inteligencia artificial describe entonces, la posibilidad de que los sistemas intenten imitar el pensamiento humano partir de aprender y utilizar las generalizaciones que usamos para tomar decisiones³. Sin embargo, la inteligencia artificial es una paradoja porque descansa en el deseo nunca alcanzado de emular los procesos cognitivos humanos⁴.

*En este reporte, se utilizará de forma genérica el término paraguas inteligencia artificial entendiéndolo como una disciplina que engloba diferentes formas de automatización y que se vale de algoritmos. Si bien se reconoce que existen diferentes conceptos como la inteligencia artificial blanda, la inteligencia artificial dura, la inteligencia artificial generativa y la inteligencia artificial general dependiendo del enfoque desde el que se analizan.

Si bien importa la claridad técnica, lo fundamental es identificar en qué medida se permite que la tecnología reemplace actividades que normalmente haría una persona como redactar una sentencia.

- **Algoritmos**

El algoritmo es clave para definir a la inteligencia artificial. El algoritmo hace referencia a una serie de instrucciones para realizar tareas⁵. Es un modelo creado por las personas que ayuda a representar la realidad⁶. Los avances que ha supuesto la tecnología han permitido que los modelos algorítmicos sean cada vez más complejos. En el caso de la inteligencia artificial, un algoritmo contiene las instrucciones que permiten a una computadora aprender del entorno y realizar tareas⁷. Una receta de cocina y el procedimiento para cambiar una llanta pinchada son ejemplos de algoritmos⁸.

- **Aprendizaje automatizado**

El aprendizaje automatizado (ML⁹ por sus siglas en inglés) permite a las máquinas aprender automáticamente utilizando patrones y deducciones en lugar de recibir instrucciones de una persona. Las técnicas de ML instruyen a las máquinas para que alcancen un resultado al proporcionar numerosas instancias de resultados correctos¹⁰. Entre más entrenan, más experiencia adquieren y pueden replicarlo de mejor forma en el futuro¹¹.

- **Big data o macrodatos**

Conjuntos de datos demasiado grandes o complejos para ser analizados por los

¹ Jamila Venturini. Inteligencia artificial más allá de la polarización. Abril, 2024. Disponible en <https://www.derechosdigitales.org/23564/inteligencia-artificial-mas-alla-de-la-polarizacion/>

² Id.

³ Jordi Nieva. Inteligencia artificial y proceso judicial, Marcial Pons, 2018, pág. 20.

⁴ Fernando Miró Linares citado en Álvaro Vizcaíno Zamora. Videovigilancia e inteligencia artificial: entre la utopía y la distopía. pág. 30.

⁵ UNESCO. Kit de herramientas global sobre IA y el Estado de derecho para el poder judicial. 2023, pág. 28. Disponible en https://unesdoc.unesco.org/ark:/48223/pf0000387331_spa

⁶ Luis Fernando García y Grecia Macías, La nueva cara de la discriminación: el impacto diferenciado de la tecnología de reconocimiento facial en grupos vulnerables en SCJN, Discriminación, Problemas Contemporáneos, Ricardo Latapie Aldana (Coord.), pág. 111.

⁷ UNESCO. Kit de herramientas, Op. Cit., pág. 28.

⁸ Juan Corvalán et. al. Inteligencia artificial: bases conceptuales para una aproximación interdisciplinar en Juan Corvalán, Tratado de Inteligencia Artificial y Derecho, T. 1.

⁹ Machine Learning.

¹⁰ UNESCO. Kit de herramientas, Op. cit., pág. 21.

¹¹ Luis Fernando García y Grecia Macías, La nueva cara de la discriminación, Op. cit., pág. 112.

programas tradicionales. La disponibilidad cada vez mayor de macrodatos, gracias al uso creciente de Internet por parte de la sociedad y a las rápidas mejoras en la potencia informática ha permitido avances significativos en la IA en los últimos 10 años¹.

- **Reconocimiento facial**

La tecnología de reconocimiento facial utiliza la identificación biométrica de una persona para mapear sus características faciales contenidas en una fotografía o en un video y las compara con la información obtenida con un banco de rostros para encontrar una coincidencia.

- **Sistemas de decisión automatizada**

Procedimientos en los que las decisiones son inicialmente delegadas, total o parcialmente, a otra persona o ente corporativo que, a su vez, ejecuta modelos automatizados de toma de decisión para el desarrollo de la acción².

4. Inteligencia artificial en labores de seguridad pública y por los sistemas de justicia penal en Europa

En septiembre de 2021, Fair Trials publicó el informe Injusticia Automática: El uso de la inteligencia artificial y de los sistemas de decisión automatizada en el sistema penal en Europa³. En este informe, Fair Trials analizó 13 sistemas de inteligencia artificial o de decisión automatizada que han sido implementados en los sistemas de justicia en Alemania, España, Holanda, Inglaterra y Gales, Italia y Ucrania. A continuación, se hará un recuento breve de algunos de los sistemas, los cuales no incluyen reconocimiento facial, y sus principales hallazgos. La exposición se divide entre sistemas de IA que fueron utilizados para labores de seguridad pública y sistemas de IA utilizado por poderes judiciales o por autoridades penitenciarias.

i. Predicción de crímenes

- *RADAR-iTE. Alemania*

En 2017, la Oficina Federal de Investigación Criminal Alemana (BKA) desarrolló RADAR-iTE, una herramienta de análisis de riesgo de violencia. Dicha herramienta fue destinada a evaluar el riesgo de "delincuentes potencialmente destructivos" de cometer actos de terrorismo islamita.

De acuerdo con la policía alemana, usando este sistema los recursos de las autoridades se dirigen de forma más específica a personas identificadas como de alto riesgo de cometer un acto violento. RADAR-iTE utiliza información policial sobre el "comportamiento observable" de una persona que es recopilada por trabajadores sociales. Se recopila tanta información como sea posible sobre eventos en la vida de la persona. A partir de ello, la herramienta clasifica a las personas como de riesgo "alto", "llamativo" o "moderado".

¹ Access Now. Human Rights in the Age of Artificial Intelligence, pág. 8.

² Patricio Velasco Fuentes y Jamila Venturini. Decisiones automatizadas en la función pública en América Latina. Una aproximación comparada a su aplicación en Brasil, Chile, Colombia y Paraguay. Derechos Digitales, Marzo de 2021., pág. 5.

³ Fair Trials, "Automating Injustice: The use of artificial intelligence & automated decision-making systems in criminal justice in Europe", Octubre de 2021. https://www.fairtrials.org/app/uploads/2021/11/Automating_Injustice.pdf

Fair Trials concluyó que por su diseño y objetivo RADAR-iTE tenía como objetivo evaluar el riesgo de personas musulmanes. De la información disponible en línea se desprende que dicha herramienta de análisis de riesgo se continúa utilizando¹.

- *ProKid. Holanda*

En Holanda, en el año 2011, ProKid fue desarrollado y utilizado por la policía para evaluar el riesgo de que niños y niñas menores de 12 años cometieran delitos. Aunque ProKid se presentó como un método de predicción de crímenes, en realidad sólo predecía la probabilidad de que un niño o una niña fuera relacionado con un delito en los registros policíacos. Los algoritmos de ProKid se alimentaron de 2 bases de datos de la policía holandesa. Dichas bases de datos contenían información de niños y niñas que habían tenido cualquier tipo de contacto con la policía ya sea como sospechosas, víctimas o como testigos de delitos. Cuando ProKid evaluaba que un niño representaba un riesgo podría haber consecuencias para él y para su familia. Incluso podría ser separado de su familia. A partir de los análisis de ProKid también se registraban antecedentes penales a sus padres.

Sin embargo, se documentó que ProKid cometía errores. Una evaluación de ProKid solicitada por el Ministerio de Seguridad y Justicia de Holanda revisó 2,444 análisis de riesgo y consideró que únicamente 1,542 eran correctos. Es decir, más de un tercio de las niñas y los niños fueron asignados con un nivel de riesgo erróneo.

No se tiene información de que el sistema continúe utilizándose si bien se han desarrollado nuevas versiones como ProKid 23.

- *NDAS. Reino Unido*

En 2016, en el Reino Unido, diferentes instituciones de policía crearon la base nacional de análisis de datos (NDAS por sus siglas en inglés) para evaluar riesgos y predecir delitos. Esta herramienta basada en aprendizaje automático realiza análisis de comportamiento y modelos predictivos para realizar predicciones sobre las posibles acciones futuras de personas en específico. A partir de dichas herramientas se busca influir en las acciones policiales de prevención de delitos.

El uso de NDAS ha llevado a que la policía detenga y revise a las personas afrodescendientes casi diez veces más que a personas blancas el Reino Unido y en Gales.

En 2021, año en el que el reporte de Fair Trials fue lanzado, la Policía de West Midlands afirmó que pretendían usar datos provenientes de oficinas públicas de salud, educación, seguridad social y de autoridades locales para alimentar las herramientas de predicción. Asimismo, al menos hasta 2021, la herramienta continuaba utilizándose.

- *HART. Reino Unido*

En el Reino Unido también se utiliza The Harm Assessment Risk Tool (HART). Se trata de un sistema automatizado adquirido por la Policía de Durham en 2017 que promete predecir si un sospechoso de delito cometerá un delito en el futuro. A partir del puntaje asignado por HART, las autoridades deciden si someterán a personas a proceso penal o la integrarán a un programa de rehabilitación.

¹ Ver The International Centre for Counter-Terrorism (ICCT) <https://www.icct.nl/sites/default/files/2024-02/Female%20Jihadis%20Facing%20Justice%20Chapter%204.pdf>

La información que alimenta HART proviene de antecedentes penales, edad de la persona, género y del lugar de residencia. A partir de una investigación realizada por Big Brother Watch, una organización de derechos humanos basada en el Reino Unido descubrió que la policía contrató a una empresa que clasifica consumidores para alimentar a HART. Big Brother Watch documentó que la empresa, llamada Mosaic incluía criterios abiertamente discriminatorios y estereotipados para clasificar a los consumidores como su ascendencia asiática, su bajo nivel de ingresos, ser de clase trabajadora o incluso su nombre. La variable proveída por Mosaic tenía un alto nivel de influencia en la clasificación de la persona como de alto, medio o bajo nivel de riesgo. La policía de Durham discontinuó el uso de Mosaic cuando fue expuesta. Sin embargo, al menos hasta 2021, HART continuaba utilizándose.

Finalmente, la precisión de HART, calculada a partir del número de individuos que realmente reincidieron fue del 62% y las predicciones de alto riesgo sólo fueron precisas solamente en un 52% de los casos.

b. Sistema de justicia penal

i. Evaluación de medidas cautelares

- *Cassandra. Ucrania*

En 2020, el Ministerio de Justicia de Ucrania anunció que se había desarrollado un sistema de evaluación que evalúa el riesgo basado en inteligencia artificial. Se dijo que Cassandra automatizaría la elaboración de reportes para dictar medidas cautelares y para el dictado de sentencias. En específico, Cassandra analiza el riesgo de reincidencia de personas que cometieron delitos. Dichos reportes sirven para apoyar las decisiones judiciales relacionadas con la libertad de personas detenidas.

El objetivo central de Cassandra es generar una predicción de probabilidades de que una persona cometa un nuevo crimen en una escala de 0 a 97. Al momento de la publicación del reporte de Fair Trials no había información pública de los datos que alimentaba a Cassandra para realizar las evaluaciones de riesgo.

ii. Libertad anticipada

- *RisCanvi. España*

En 2009, en España se implementó RisCanvi, un sistema de evaluación de riesgo basado en algoritmos que se utiliza para apoyar la decisión de si personas detenidas pueden recibir tratamiento de libertad anticipada. RisCanvi es utilizado por el sistema penitenciario, pero vale la pena revisar su funcionamiento. En 2018 se realizaron 15 mil evaluaciones con RisCanvi.

Los factores de riesgo utilizados por RisCanvi incluyen antecedentes penales, la edad de la persona cuando cometió el delito, la duración de la pena, el tiempo que lleva en prisión, intentos de escaparse, así como reportes disciplinarios. RisCanvi también usa información sobre el nivel educativo de la persona. Por ejemplo, habilidad mental baja, daño mental severo o intentos de autolesionarse. También utiliza antecedentes penales de la familia, si tiene amigos criminales o antisociales, la promiscuidad sexual y el abuso de drogas alcohol.

La decisión sobre la libertad anticipada no es tomada al 100% por el sistema porque se acompaña de entrevistas realizadas por personal penitenciario. Los resultados de las

entrevistas son introducidos a RisCanvi y a partir de eso, el sistema predice si el riesgo es bajo, medio o alto. Un Comité decide si acepta la decisión de RisCanvi la desecha con una justificación.

Sin embargo, de acuerdo con el informe de Fair Trials, es muy difícil para una persona privada de la libertad apelar la decisión de RisCanvi porque los jueces que resuelven consideran que la decisión es tomada por las autoridades penitenciaria a pesar de que es altamente influenciada por RisCanvi.

A pesar de las críticas, el sistema se sigue utilizando hasta la fecha¹.

5. Conclusiones preliminares

El uso de diferentes sistemas de decisión automatizada en temas relacionados con la seguridad y la justicia en Europa ha develado la criminalización y la discriminación de personas por su origen étnico, por sus ingresos o incluso por su lugar de residencia.

Asimismo, las evaluaciones realizadas han demostrado que en general, los sistemas cometen errores tan graves que ponen en duda su efectividad para apoyar las labores de combate al crimen y de decisiones judiciales relevantes.

Además, debido a la falta de transparencia y acceso a la información sobre las bases de datos que alimentan los sistemas, así como sobre la forma en que se adopta la decisión, las personas no tienen posibilidades de defenderse una vez que son analizadas por algún sistema de decisión automatizada.

Varios de los sistemas basados en inteligencia artificial usados en Europa en tareas relacionadas con la seguridad pública y con la justicia han sido reformados a partir de las investigaciones realizadas por la sociedad civil. Resulta entonces fundamental continuar el acercamiento crítico a su evolución y funcionamiento, evitando caer en el mito del tecnosolucionismo.

La experiencia de Europa enseña que el uso de la automatización y de los algoritmos puede profundizar la discriminación de poblaciones en situación de vulnerabilidad si no se cuida la información que alimentará los sistemas. Por ello, se requiere regulación, transparencia, rendición de cuentas y la posibilidad de revisar judicialmente las decisiones adoptadas por sistemas de inteligencia artificial.

6. Inteligencia artificial en labores de seguridad pública y en la administración de la justicia penal en América Latina

Centrándose en América Latina, este apartado sigue la lógica del anterior para presentar experiencias de uso de sistemas basados en inteligencia artificial tanto en labores de seguridad pública como en las tareas de la administración de justicia y de la administración de la justicia penal en América Latina. Esta división no es homogénea, pero, por regla general, los sistemas de IA para labores de seguridad pública son

¹ El País. El algoritmo que evalúa la peligrosidad de los presos catalanes funciona de forma "azarosa" Julio 2024, Ver <https://elpais.com/tecnologia/2024-07-03/el-algoritmo-que-evalua-la-peligrosidad-de-los-presos-catalanes-funciona-de-forma-azarosa.html>

adquiridos e implementados por cuerpos policíacos, secretarías de seguridad o ministerios de seguridad. Mientras tanto, los sistemas de AI para la administración de justicia penal son adquiridos y utilizados por poderes judiciales.

a. Seguridad pública

En primer lugar, se expone el uso de herramientas de inteligencia artificial para reconocimiento facial y para predicción de crímenes. En materia de seguridad, la tecnología que más ha proliferado es la de reconocimiento facial. Brasil es el país donde se reporta un mayor uso general de tecnologías de vigilancia¹. Sin embargo, también se revisan otros países.

i. Reconocimiento facial

- *Reconocimiento facial (Oi). Brasil.*

En el estudio realizado por la organización de derechos humanos O Panóptico, se reportó que la policía militar de Río de Janeiro implementó un programa piloto de reconocimiento facial desde el 2019 como una solución para hacer frente a la criminalidad. El programa fue implementado con la intervención de la empresa Oi².

Sin embargo, de acuerdo a documentos internos de la policía militar, el sistema de reconocimiento facial fue diseñado para apoyar el alegato de inocencia de la policía frente a posibles denuncias por su actuación ilegal. En otras palabras, la evidencia producida por la tecnología de reconocimiento facial sólo sería usada si coincidía con la versión policiaca de la detención³.

Las cámaras fueron colocadas en puntos estratégicos de la ciudad de Río de Janeiro. De acuerdo con los documentos analizados por O Panóptico, existe un sesgo discriminatorio racial hacia personas afrodescendientes, así como de género, dando un resultado de falsos positivos al momento de su detención⁴.

- *Policía de proximidad. México*

El análisis de la inteligencia artificial en México en tareas de seguridad requiere revisar el uso de cámaras de videovigilancia (CVV). En 2018, el Instituto Nacional de Estadística, Geografía e Informática (INEGI) dio a conocer que en México había 53,949 CVV⁵. Por su parte, el gobierno de la Ciudad de México informó en el año 2020 que contaba con 15,310 CVV⁶.

En 2016, una investigación académica documentó que las personas que operaban los sistemas de videovigilancia seguían a trabajadores de la construcción que se desplazaban en zonas residenciales del Estado de México. También seguían a personas de la comunidad LGBTTTQ+ en una plaza comercial de la Ciudad de México. Quienes operaban las cámaras valoraron que los trabajadores de la construcción podían entrar a robar a algún edificio porque no era habitual ver a personas caminando en esas zonas

¹ Instituto Igarapé, Implementação De Tecnologias De Vigilância No Brasil e na América Latina, Artigo Estratégico 58, Novembro de 2022, pág. 2. Disponible en <https://igarape.org.br/wp-content/uploads/2022/12/Implementacao-de-tecnologias-de-vigilancia-no-brasil-e-na-america-latina.pdf>

² Nunes, Pablo A Rio of cameras with selective eyes: the use of facial recognition by the Rio de Janeiro state police / Pablo Nunes, Mariah Rafaela Silva, Samuel R. de Oliveira. – Rio de Janeiro : CESeC, 2022. pág. 7. Disponible en <https://drive.google.com/file/d/1ofwajPJR9EAYQ48cuMG8qLsYCf0VEuDu/view>

³ Nunes, Pablo A Rio of cameras with selective eyes: the use of facial recognition by the Rio de Janeiro state police, Op. cit., pág. 16.

⁴ Id.

⁵ Álvaro Vizcaíno Zamora. Videovigilancia e inteligencia artificial: entre la utopía y la distopía., pág. 10.

⁶ Ibid, págs. 10 y 11.

residenciales de la ciudad. En el caso de las personas de la diversidad sexual, el objetivo de seguirlas era monitorear si se besaban o abrazaban. Si les detectaban, los cuerpos de seguridad les pedían comportarse de otra manera o les invitaban a salir del lugar¹.

Las cámaras de videovigilancia no solamente son operadas por personas, también incorporan la tecnología de reconocimiento facial. De este modo, la inteligencia artificial aplicada al reconocimiento facial permite a las policías tomar decisiones relacionadas con la seguridad pública, es decir, detener a personas. En México, se han reportado detenciones basadas en el uso de tecnología de reconocimiento facial en Aguascalientes², y en Coahuila³

Frente a ello, se ha documentado que la tecnología de reconocimiento facial reproduce los sesgos de las personas que etiquetan los datos de entrenamiento de esta tecnología⁴. Además, estos sistemas cometen errores debido a factores externos como las condiciones de luz, la resolución de las imágenes o el ángulo del rostro⁵.

También en México, en el año 2020, la Dirección Seguridad Ciudadana del municipio de Nezahualcóyotl, Microsoft y Analytikis crearon el Modelo de Policía de Proximidad con el objetivo de predecir dónde puede ocurrir un delito. Para ello, desarrollaron un programa de inteligencia artificial con capacidades analíticas y predictivas que utilizan los datos recolectados a través de sistemas de vigilancia y servicio público con el objetivo de proporcionar inteligencia e información que facilite la toma de decisiones tácticas y estratégicas. De acuerdo con CodinRights, ello puede presentar problemas de estigmatización, discriminación y criminalización de las personas más vulnerables⁶.

- *Calle Segura. Chile.*

En marzo de 2019, el entonces presidente de Chile, Sebastián Piñera, dio a conocer el "Sistema de televigilancia móvil como parte del plan "Calle Segura". Dicho sistema fue anunciado para utilizar aeronaves no tripuladas, es decir, drones, en el combate a la delincuencia. Los drones estarían equipados con tecnología de reconocimiento facial. Según la información, se planeaba implementar el monitoreo en la Región Metropolitana y después extenderla a todo el país.

Frente a ello, un grupo de organizaciones de derechos humanos se pronunciaron en contra del uso de esta tecnología. Algunas de las razones para rechazar el uso de tecnología de reconocimiento facial fueron: es altamente falible, facilita la institucionalización de sesgos asociados a la clase social y al color de piel y carece de mecanismos de transparencia y rendición de cuentas⁷.

- *Reconocimiento facial. El Salvador*

En El Salvador, el gobierno de Nayib Bukele anunció la adquisición de cámaras de vigilancia con capacidad para realizar reconocimiento facial como parte de su política de seguridad. Las cámaras fueron adquiridas con recursos provenientes de un préstamo del

¹ Ibid., págs. 14 y 15.

² H. Ayuntamiento de Aguascalientes 2019-2011. Boletín no.1161 "Nuevas Cámaras de Seguridad Arrojan Resultados positivos". Consulta el día 7 de julio de 2021 en: <https://www.ags.gob.mx/cont.aspx?p=6253> citado en Luis García y Grecia Macías. La nueva cara de la discriminación: el impacto diferenciado de la tecnología de reconocimiento facial en grupos vulnerables en SCJN. Discriminación, Problemas Contemporáneos, Ricardo Latapie, Coord., 2022, México. pág. 124

³ Luis García y Grecia Macías. La nueva cara de la discriminación: el impacto diferenciado de la tecnología de reconocimiento facial en grupos vulnerables en SCJN. Discriminación, Problemas Contemporáneos, Ricardo Latapie, Coord., 2022, México. pág. 125.

⁴ Ibid., Pág. 139.

⁵ Ibid., Pág. 124.

⁶ Disponible en <https://notmy.ai/es/project-item/modelo-de-policia-de-proximidad-es/>

⁷ Datos protegidos, Pronunciamiento "Estamos en contra de la vigilancia masiva en los espacios públicos del "Sistema de televigilancia móvil", Abril 8-2019. Dicha publicación fue suscrita por otras ONG'S. <https://datosprotegidos.org/estamos-en-contra-de-la-vigilancia-masiva/>

Banco Centroamericano de Integración Económica. BCIE¹.

Servicio Social Pasionista, una organización de derechos humanos salvadoreña, se mostró preocupada por la opacidad con la que está adquiriéndose y operando este tipo de tecnología. Además, debido a que la política de seguridad implementada en El Salvador se ha caracterizado por cometer graves abusos a los derechos humanos, temen que el uso de estos sistemas sea contraproducente, especialmente contra poblaciones típicamente estigmatizadas y discriminadas².

ii. Predicción de delitos

En Chile, Colombia y Uruguay se han implementado o se han realizado pruebas piloto de sistemas de predicción de delitos. En algunos casos se anunció el uso de la tecnología pero no existe información certera sobre si los gobiernos están utilizándola.

- *Sistema Predictivo del Delito Urbano. Chile*

El Sistema Predictivo del Delito Urbano fue desarrollado en 2017 por la Universidad de Chile³ junto con el Departamento de Análisis Criminal (DAC) de Carabineros de Chile. Fue implementado con el objetivo de predecir zonas de mayor riesgo de ocurrencia de delitos para dirigir de forma eficiente el patrullaje policial⁴.

De acuerdo con el informe técnico se buscaba desarrollar un sistema predictivo delictual basado en algoritmos matemáticos con la finalidad de realizar patrullajes focalizados y de reforzar "la eficacia del sistema de persecución penal".

El sistema fue analizado por la investigadora Josefina Buschmann para Derechos Digitales América Latina. Algunas de las conclusiones del análisis son las siguientes.

No existen los datos neutrales u objetivos. Se utilizan datos policiales respecto de los cuales deben analizarse su contexto de producción. El contexto de producción puede ser detenciones injustas o arbitrarias o denuncias sobre/subestimadas. Dicho contexto puede sesgar la base de datos y reproducir prácticas discriminatorias que lleven a intensificar la vigilancia y los controles de identidad preventivos en zonas determinadas⁵.

Frente a esos problemas, no existe un protocolo de auditoría externa y de carácter civil de las prácticas de recolección de datos de la policía⁶.

Uno de los puntos centrales de la evaluación afirma que el "problema de integrar datos policiales en un sistema [...] automatizado [...] es que los rearticula como datos con validez científica"⁷.

Los carabineros, –la policía chilena– no consideraron que el modelo puede impactar los derechos de las personas porque no utiliza datos personales o permite identificar lugares exactos. Sin embargo, Josefina Buschmann apunta las zonas de vigilancia creadas a partir del sistema puede implicar cambios en la conducta de la policía. Estos cambios

¹ SSPAS, Uso de tecnología en las políticas de seguridad en El Salvador, los riesgos en los derechos humanos. 2021, págs. 23, 25 y 32.

² Ibid., ver conclusiones.

³ En específico, por el Centro de Análisis y Modelamiento en Seguridad (CEAMOS).

⁴ Josefina Buschmann, Sistema Predictivo del Delito Urbano; Producción algorítmica de Zonas de Vigilancia y Control en la Ciudad, Derechos Digitales, Inteligencia Artificial e Inclusión en América Latina. Diciembre 2021. Pág. 6, Disponible en https://www.derechosdigitales.org/wp-content/uploads/02_Informe-Chile-Sistema-Predictivo-del-Delito-Urbano_ES_28042022.pdf

⁵ Ibid., págs. 41 y 42.

⁶ Ibid., pág. 42.

⁷ Id.

pueden incluir el hostigamiento a personas que no han cometido un delito pero cuyo comportamiento es considerado sospechoso¹.

- *Estudio empírico de identificación de comportamiento criminal. Chile*

En 2018, la Pontificia Universidad Católica de Chile publicó una investigación² relacionada con el uso de herramientas automatizadas para analizar y predecir el comportamiento delictivo y ayudar a la policía a reducir el crimen. El estudio intentó predecir el comportamiento criminal de personas que hubieren sido detenidas previamente³. El modelo de predicción propuesto no fue aplicado.

La investigación analizó casi 800 mil detenciones realizadas por las policías entre 2009 y 2018 y la información sociodemográfica sobre los victimarios. A partir de ello, estimó las posibilidades de reincidencia de una persona. Algunos elementos que fueron tomados en cuenta fueron: detenciones previas y antecedentes personales, incluyendo el número de detenciones de familiares.

El 37% de las predicciones realizadas por el modelo de que una persona volvería a ser detenido fueron erradas⁴. A las predicciones erradas se les calificó como falsos positivos y se reconoció que implican un alto costo social. La investigación que el delito es simplemente difícil de predecir. Además, afirma que la predicción se puede mejorar si se agregan atributos como la historia de las detenciones o indicadores biográficos como educación o los bienes de las personas⁵.

El estudio hace preguntas clave éticas y legales. Por ejemplo: ¿hasta qué punto el Estado puede utilizar la información personal para fines que no han sido informados ni consentidos por los ciudadanos? ¿Se puede utilizar el registro de detenciones anteriores u otra información similar como base para realizar acciones de prevención e investigación criminal?

De este modo, la investigación reconoce que este tipo de sistemas pueden alentar la detención de personas por sus antecedentes personales o familiares y no por conductas que realmente hayan cometido. Además, abre la posibilidad de afectar derechos de personas en situación de vulnerabilidad.

- *Colombia Policías. Colombia*

El gobierno de Colombia dio a conocer en el año 2019⁶, el proyecto “Colombia Policías”, cuyo objetivo es anticipar el crimen en Bogotá, prediciendo cuándo, dónde, quien y porqué del crimen a través de un programa matemático predictivo que describe los homicidios, los delitos contra la propiedad que impliquen fuerza o amenaza de fuerza, las lesiones personales y la dinámica detrás de la percepción pública de seguridad. El programa fue desarrollado por la Secretaría Distrital de Seguridad, Convivencia y Justicia.

De acuerdo con el mapeo de CodinRights⁷, preocupa la posible estigmatización, discriminación y criminalización de las personas más vulnerables⁸.

¹ Josefina, Buschmann, Sistema Predictivo del Delito Urbano, Op. cit., pág. 47.

² Van 'T Wout et. al. Big data para la identificación de comportamiento criminal, en Propuestas para Chile, Pontificia Universidad Católica de Chile, 2018, Disponible en: https://politicaspUBLICAS.uc.cl/content/uploads/2019/03/Libro-completo-en-PDF_final-6.pdf

³ Ibid., pág. 51.

⁴ Ibid., pág. 65.

⁵ Ibid., pág. 66.

⁶ El Espectador, “El matemático que quiere predecir los crímenes de Bogotá”, 24 jun 2019. <https://www.elespectador.com/ciencia/el-matematico-que-quiere-predecir-los-crmenes-de-bogota-article-867606/>

⁷ CodinRights <https://notmy.ai/es/mapeo-de-proyectos/>

⁸ Disponible en <https://notmy.ai/es/project-item/modelo-de-policia-de-proximidad-es/>

- *Predicción del delito en ciudades intermedias. Colombia*

Una investigación publicada en septiembre de 2022 analizó el uso de la tecnología basada en inteligencia artificial para predecir el delito en una ciudad intermedia, Bucaramanga, ubicada en el Departamento de Santander en Colombia. La investigación calculó la capacidad predictiva del delito en una ciudad respecto de la que se tienen pocos datos y pocas capacidades en comparación con ciudades principales como Bogotá¹.

La investigación desarrolló un modelo piloto de predicción del delito que no fue aplicado. Su fuente de datos fue el Sistema de Información Estadísticos, Delincuencial, Contravencional y Operativo de la Policía Nacional (SIEDCO). El nivel de precisión obtenido en la predicción fue de entre el 50% y 60%, lo que significa que cerca de la mitad de los pronósticos que efectuó el modelo eran verdaderos delitos².

La investigación reconoce la existencia de sesgos y discriminación en los modelos predictivos e invita a considerar estos factores antes de su implementación³.

En las conclusiones, se señala que los modelos de predicción del crimen pueden tener efectos discriminatorios sobre determinados grupos poblacionales o en zonas de la ciudad. Por ello, recomienda implementar los modelos de predicción en paralelo con otros esfuerzos⁴.

- *PredPol. Uruguay*

En el año 2013, el Ministerio del Interior de Uruguay adquirió el software PredPol, el cual surgió de un proyecto de investigación conjunto del Departamento de Policía de los Ángeles y de la Universidad de California⁵.

La herramienta prometía predecir dónde ocurrirá un delito. (en espacios de 150 metros cuadrados). PredPol fue aplicado entre 2014 y 2017⁶.

PredPol realizó predicciones a partir del sistema que aglutina la información de seguridad pública en todo Uruguay.

El análisis del funcionamiento de la herramienta concluyó que no todos los ilícitos son denunciados, y las bases de la policía, como es de esperar, se construyen a partir de lo reportado por las personas. Ello implica, por ejemplo, que algunos delitos se reportan más que otros.

Asimismo, se identificó que la presencia policial en una zona puede incrementar las probabilidades de que se identifique o reporte un delito

Cumplido el proceso de evaluación, el Ministerio del Interior observó que las áreas donde se operó con PredPol no mostraban mejores resultados que aquellas donde se utilizó el sistema estático desarrollado por sus propios técnicos. Por esa razón, se optó por discontinuar el uso de PredPol⁷.

¹ Juan-David Galvez-Ferreira et. al. Prediciendo el crimen en ciudades intermedias: un modelo de "machine learning" en Bucaramanga, Colombia

² Ibid., pág. 94.

³ Ibid., pág. 85.

⁴ Ibid., págs. 95 y 96.

⁵ WWF. Algoritmos e inteligencia artificial en América Latina. Un estudio de implementaciones por parte de Gobiernos de Argentina y Uruguay, 2018, pág. 27.

⁶ Id.

⁷ WWF. Algoritmos e inteligencia artificial en América Latina, Op. cit., pág. 29.

b. Administración de la justicia y justicia penal

i. Valoración de pruebas y sentencias. México

En México, no existe registro de que la utilización de inteligencia artificial en el sistema penal. Sin embargo, esa posibilidad ha sido discutida. Por ejemplo, Juliana Vivar publicó un artículo¹ en la revista del Instituto Nacional de Ciencias Penales (INACIPE) en el que analizó teóricamente la posibilidad de utilizar modelos de predicción en sentencias penales a partir del marco legal mexicano.

En el artículo, Juliana Vivar analiza las características que distinguen al ser humano de la máquina como la libertad y la autonomía así como las características que le son comunes como el error y los sesgos.

ii. Medidas cautelares

- *PRiSMA. Colombia*

En el año 2016, la Fiscalía General de la Nación en Colombia inició un programa operado mediante inteligencia artificial, denominado PRiSMA (Perfil de Riesgo de Reincidencia para Solicitud de Medidas de Aseguramiento), con el objetivo de determinar el riesgo de reincidencia de personas acusadas de delitos para apoyar la decisión de la Fiscalía sobre si solicitar o no la detención preventiva.

PRiSMA funciona a partir de un algoritmo que mediante aprendizaje automatizado procesa información de una base de datos. La base de datos fue dotada de información por la Policía Nacional y por la Fiscalía General de la Nación². A partir de ello, PRiSMA apoyaba a predecir la probabilidad de que una persona reincidiera en la comisión de un delito. Si el sistema predice que la persona representa un nivel de riesgo alto, la Fiscalía solicita una medida cautelar más restrictiva para la libertad con el fin de evitar que cometa un delito. Si el sistema predice un riesgo bajo, se solicitan medidas no restrictivas de la libertad.

La base de datos que alimenta a PRiSMA tiene información de seis millones de individuos con registros de antecedentes penales³. Usando aprendizaje automatizado se determinan patrones de comportamiento asociados a eventos delictivos, actuales y previos, registros judiciales, contravenciones y penitenciarios y se establece un nivel de riesgo relacionado con la comisión de delitos contra el patrimonio, delitos violentos y otros delitos, en un periodo de dos años posterior a la imputación.

La Fiscalía General de la Nación⁴ esperaba poder no sólo disminuir la reincidencia criminal sino también reducir el uso de la prisión preventiva. Sin embargo, también podría preocupar que las garantías del debido proceso penal como el derecho a la defensa, la presunción de inocencia y la imparcialidad judicial puedan verse afectadas como parte del

¹ Juliana Vivar. Sentencia y predicción algorítmica penal. Herramienta o suplencia humana en INACIPE, Ciencias Penales e inteligencia artificial, 2021, Disponible en <https://revistaciencias.inacipe.gob.mx/index.php/02/article/view/437/342>

² Carmen Eloísa Ruiz López, Valentina del Sol Salazar y Humberto J. Sierra. Sistemas Operados mediante Inteligencia Artificial (IA) y debido proceso penal. Perspectiva de aplicación en Colombia en Disrupción tecnológica, transformación digital y sociedad, Tomo III, Derecho, Innovación y tecnología: Fundamentos para el mundo digital. Diego Acosta González (y otros): Juan Carlos Henao, Daniel Castaño (eds).- Bogotá, Universidad Externado de Colombia, 202, pág. 841. Disponible en <https://bdigital.uexternado.edu.co/server/api/core/bitstreams/97830f4d-62c2-4d3c-b3cc-fea2145e11c8/content>

³ V. SALAZAR. Estrategias para la racionalización de la detención preventiva en Colombia: un análisis del instrumento PRiSMA, Colección Tesis de grado No. 107, Bogotá, Universidad Externado de Colombia, 2020, Título 4.1; crf. También, Centro de Estudios sobre Seguridad y Drogas (CESED). "Prisma: el programa de la Fiscalía para predecir la reincidencia criminal", Facultad de Economía de la Universidad de los Andes, disponible en [<https://cesed.uniandes.edu.co/prisma-el-programa-de-la-fiscalia-para-predecir-la-reincidencia-criminal/>] citado en Disrupción tecnológica, transformación digital y sociedad

⁴ Fiscalía General de la Nación; Herramienta PRiSMA, Perfil de Riesgo de Reincidencia para la Solicitud de Medidas de Aseguramiento, Dirección de Políticas Públicas y Estrategia. <https://www.fiscalia.gov.co/colombia/wp-content/uploads/Perfil-de-riesgo-de-reincidencia-para-solicitudes-de-medida-de-aseguramiento.pdf>

riesgo del uso de la inteligencia artificial¹.

Codin Rights consideró que PRiSMA puede incurrir en algunos sesgos como estigmatización, discriminación y criminalización de las personas más vulnerables².

No se tiene información de que PRiSMA se continúe utilizando.

iii. Valoración de prueba y dictado de sentencias

- *Projeto Sócrates Proceso penal. Brasil*

En Brasil, en el año 2019 se anunció el “Projeto Sócrates” el cual mediante inteligencia artificial que busca reducir en un 25% el tiempo para emitir sentencias de apelación³. El sistema analiza las apelaciones que recibe el Tribunal Superior de Justicia de Brasil a partir de 300,000 casos resueltos y agrupa los casos que son similares para decidirlos juntos⁴.

De acuerdo con el ministro de justicia, Ricardo Villas, se busca implementar que mediante este sistema se produzcan proyectos de decisiones automatizadas a partir del análisis de decisiones previas⁵.

Sin embargo, de acuerdo con Codin Rights, preocupa el posible riesgo de automatizar las injusticias sistemáticas producidas históricamente por el sistema judicial brasileño⁶.

iv. Gestión judicial

- *PROMETEA. Argentina*

En 2019, la Fiscalía General Adjunta en lo Contencioso Administrativo y Tributario del Ministerio Público Fiscal de Buenos Aires, Argentina adoptó PROMETEA. Este sistema incorpora inteligencia artificial para automatizar la elaboración de dictámenes jurídicos basándose en casos análogos respecto de los cuales existen precedentes judiciales reiterados⁷.

De acuerdo con sus desarrolladores, PROMETEA implicó reducir de 90 minutos a 1 minuto la resolución de un pliego de contrataciones⁸. Es decir, es un sistema que representa una eficiencia extraordinaria. La ganancia de tiempos de resolución de los casos permitió al personal dedicar más tiempo a los casos más complejos.

Una publicación del Banco Interamericano de Desarrollo, PROMETEA ha significado "una mejora significativa en la eficiencia global de la institución [la Fiscalía General] ?

¹ Disrupción tecnológica, transformación digital y sociedad, Tomo III, Derecho, Innovación y tecnología: Fundamentos para el mundo digital. Diego Acosta González (y otros): Juan Carlos Henao, Daniel Castaño (eds).- Bogotá, Universidad Externado de Colombia, 2021.

² Codin Rights, Perfil de Riesgo de Reincidencia para la Solicitud de Medidas de Aseguramiento-Prisma (Colombia). Disponible en <https://notmy.ai/project-item/prisma-en/>

³ Projeto-piloto do Sócrates, programa de Inteligencia artificial do STJ, é esperado para agosto. <https://www.migalhas.com.br/quentes/299820/projeto-piloto-do-socrates--programa-de-inteligencia-artificial-do-stj--e-esperado-para-agosto>

⁴ Leonie Wittenstein y Felipe Franco. [Uso de la inteligencia artificial y los sistemas algorítmicos en el sistema judicial](https://www.kas.de/documents/271408/16552318/Uso+de+la+Inteligencia+Artificial+y+los+sis+temas+algorítmicos+en+el+sistema+judicial+%281%29.pdf/7cd2c325-6b7f-475e-2a8b-59b5544823cb?version=1.0&t=1680036938780). KAS, Pág. 8. Disponible en <https://www.kas.de/documents/271408/16552318/Uso+de+la+Inteligencia+Artificial+y+los+sis+temas+algorítmicos+en+el+sistema+judicial+%281%29.pdf/7cd2c325-6b7f-475e-2a8b-59b5544823cb?version=1.0&t=1680036938780>

⁵ Codin Rights, Projeto Sócrates (Brasil). Disponible en <https://notmy.ai/es/project-item/projeto-socrates/>

⁶ Id.

⁷ Elsa Estevez, Sebastián Linares y Pablo Fillottrani. PROMETEA. Transformando la administración de justicia con herramientas de inteligencia artificial, BID, 2020. Disponible en <https://publications.iadb.org/es/publications/spanish/viewer/PROMETEA-Transformando-la-administracion-de-justicia-con-herramientas-de-inteligencia-artificial.pdf>

⁸ Ibid., pág. 22.

⁹ Pág. 88

La misma publicación reconoce que "el mayor riesgo de la aplicación de IA en procesos de toma de decisiones es que amplifiquen los sesgos existentes en lugar de contribuir a remediarlos". Sin embargo, en el caso de PROMETEA, dichos riesgos fueron atendidos mediante controles basados en derechos humanos².

- *PretorIA. Colombia*

La Corte Constitucional de Colombia está desarrollando PretorIA, una herramienta basada en inteligencia artificial que busca eficientar la selección de casos de tutela judicial³. En Colombia, la Corte Constitucional fija su jurisprudencia a partir de la decisión de las tutelas⁴. PretorIA, básicamente ayudará a la Corte Constitucional a preseleccionar las tutelas que serán revisadas. La Corte Constitucional recibe más de 600,000 tutelas por año⁵.

PretorIA se inspiró en PROMETEA, la herramienta usada por la Fiscalía General de Buenos Aires mencionada en este mismo reporte.

Saavedra y Upegui no vislumbran riesgos tangibles para los derechos de las personas por el uso de PretorIA debido a que se ha anunciado que automatizará parte de las actividades en el preprocesamiento de la información, pero quien tomará la decisión es una persona⁶.

7. Inteligencia artificial y desafíos emergentes en América Latina frente al debido proceso

Las investigaciones realizadas por organizaciones de derechos humanos en diferentes partes mundo han coincidido en que el uso de sistemas basados en inteligencia artificial en labores de seguridad pública y en los sistemas de justicia, representan un riesgo para los derechos humanos. La IA en el sistema penal puede profundizar la discriminación y afectar el derecho a la privacidad; además, tiene implicaciones éticas.

Sin embargo, también es necesario revisar las posibles afectaciones de la utilización de herramientas basadas en inteligencia artificial en el sistema penal para las garantías del debido proceso, como la presunción de inocencia, el derecho de defensa, el derecho a recurrir el fallo y la independencia e imparcialidad judicial. También debe tomarse en cuenta el derecho a la libertad personal⁷.

Por ejemplo, las herramientas de reconocimiento facial son utilizadas para identificar personas con fines de investigación penal. Cuando dichas herramientas cometen errores afectan la presunción de inocencia y el debido proceso y dada la dificultad de cuestionar la decisión tomada por un sistema automatizado el derecho a la defensa se torna ilusorio. Por esa razón, es necesario identificar los riesgos para los derechos humanos que las tecnologías basadas en inteligencia artificial generan. Esta identificación permitirá regular

¹ Ibid., pág. 89

² Juan G. Corvalán, Enzo Maria Le Fevre Cervini, CERIDAP, 2020. Disponible en <https://ceridap.eu/prometea-experience-using-ai-to-optimize-public-institutions/?lng=en#:~:text=Prometea%20is%20a%20predictive%20artificial,exponential%20optimizer%20of%20bureaucratic%20processes>.

³ Víctor Saavedra y Juan Carlos Upegui. PretorIA y la automatización del procesamiento de causas de derechos humanos. Derechos Digitales América Latina y Dejusticia. Pág. 18. Disponible en https://www.dejusticia.org/wp-content/uploads/2021/04/CPC_informe_Colombia.pdf

⁴ Id.

⁵ Ibid., pág. 22.

⁶ Ibid., pág. 53.

⁷ Fair Trials señaló que el uso de sistemas basados en inteligencia artificial en el sistema penal vulnera el derecho al juicio justo y la presunción de inocencia. Fair Trials, Injusticia automatizada. Op. cit., pág. 4.

su uso cuando proceda o prohibirlos.

Ahora bien, existen labores relacionadas con el sistema de justicia que son mecánicas como la búsqueda de precedentes judiciales o la generación de estadísticas y que no inciden en la determinación de derechos. En esos casos, la inteligencia artificial puede contribuir a fines nobles como mejorar las capacidades de investigación, la eficiencia y la reducción de tiempos de espera y mayor previsibilidad de las decisiones, entre otras.

En ese entendido, el acercamiento a la IA en el sistema de justicia debe ser equilibrado para no condenar su utilización a priori pero tampoco incorporarla sin reparar en sus riesgos. Cuando se condena la IA sin un diagnóstico se pueden desaprovechar sus beneficios. Cuando se acepta la IA sin revisar sus riesgos se le pueden atribuir beneficios o capacidades que no tiene.

a. Libertad personal

En los sistemas legales, cualquier afectación al derecho a la libertad realizada por un policía a una persona debe estar justificada objetivamente. Por ejemplo, cuando la policía observa a alguien cometiendo un delito. Por esa razón, el artículo 9 del Pacto Internacional de Derechos Civiles y Políticos (PIDCP o el Pacto) exige que existan causas justificadas y un procedimiento establecido en la ley para detener a una persona. La Convención Americana sobre Derechos Humanos (CADH) coincide con el Pacto. Señala que es posible restringir la libertad únicamente cuando existen causas y condiciones fijadas previamente por las Constituciones o por las leyes dictadas conforme a ellas y de acuerdo con procedimientos objetivamente definidos¹.

Reconociendo que es legítimo que la policía realice funciones de prevención del delito, las citadas normas internacionales están destinadas a proteger a las personas frente a injerencias arbitrarias o ilegales del Estado que pongan en riesgo su libertad o incluso su vida. Dichas normas son aplicables también cuando las autoridades se apoyan de la tecnología para realizar funciones de seguridad.

De este modo, realizar seguimiento de personas y revisiones aleatorias basadas en herramientas de inteligencia artificial, afecta entre otros derechos, la libertad personal. En primer lugar porque los sistemas reproducen sesgos discriminatorios que afectan especialmente a grupos sociales desventajados. Ello conlleva una restricción arbitraria a la libertad (art. 7.3 de la CADH). En segundo lugar porque los sistemas cometen errores en un porcentaje tan alto que permite a la policía seguir y revisar a personas sin una causa justificada lo cual también representa una restricción arbitraria de la libertad (Art. 7.3 de la CADH). En tercer lugar, porque dada la falta de información relacionada con la forma en la que operan las herramientas basadas en inteligencia artificial, la persona no conoce las razones que llevaron a la policía a revisarla (Art. 7.4 de la CADH). Esto último vulnera, a su vez, el derecho de toda persona a que se controle judicialmente la detención y a impugnar su legalidad. (Art. 7.5 y 7.6 de la CADH).

b. Debido proceso

El uso de la inteligencia artificial en el sistema penal requiere reflexionar sobre la aplicación de las garantías del debido proceso. En este reporte no se hará un estudio exhaustivo sobre todas las garantías del debido proceso afectadas pero sí una aproximación inicial que permita establecer que cuando se utilizan herramientas basadas en inteligencia artificial en el proceso penal los derechos humanos deben estar en el

¹ Corte IDH. Caso Fernández Prieto y Tumbeiro Vs. Argentina. Fondo y Reparaciones. Sentencia de 1 de septiembre de 2020. Serie C No. 411, párr. 66.

centro¹.

El respeto al debido proceso o juicio justo es crucial para la construcción de una decisión judicial que afecta los derechos de las personas. Garantizar el juicio justo dota de validez y de legitimidad al proceso de determinación judicial de derechos². Al final de cuentas, someter a proceso penal a una persona implica privarla de derechos tan valiosos como su libertad o sus bienes. Así, la incorporación de la inteligencia artificial al proceso penal no exenta a las autoridades de respetar este derecho de primer orden.

Por ejemplo, cuando una fiscalía o un juzgado se apoya total o parcialmente en un sistema de inteligencia artificial para decidir qué medida cautelar solicitar u ordenar, para valorar pruebas o para pronunciarse sobre un pedido de libertad anticipada afecta el debido proceso. A continuación, revisaremos algunas garantías del debido proceso que pueden ser afectadas por el uso de la inteligencia artificial.

c. Presunción de inocencia

El derecho a la presunción de inocencia es una piedra angular del debido proceso. Implica que no se presume la culpabilidad de una persona acusada de delito mientras no se haya demostrado la acusación fuera de toda duda razonable. Por esa razón, la persona acusada debe gozar del beneficio de la duda y ser tratada como inocente³.

Este derecho puede ser vulnerado desde diferentes perspectivas por sistemas de inteligencia artificial. Por ejemplo, cuando un sistema predice que una persona cometerá un delito y a partir de ello se toman decisiones importantes como negarle la posibilidad de ser juzgada en libertad se afecta la presunción de inocencia. Es decir, la persona es considerada con un nivel de riesgo determinado por coincidir con un perfil previamente diseñado⁴.

No es necesario que la persona sea detenida para afectar su derecho a la presunción de inocencia. Es suficiente que la policía realice revisiones supuestamente aleatorias a partir de decisiones que descansan en tecnologías que incorporan inteligencia artificial para causar una afectación.

d. Derecho de defensa e igualdad

Toda persona acusada de un delito tiene derecho a ser informada en un idioma que comprenda y en forma detallada, de la acusación formulada contra ella (art. 14.2.a del PIDCP). También tiene derecho a disponer del tiempo y de los medios apropiados para preparar su defensa (art. 14.2.b del PIDCP) y a interrogar a los testigos que la acusan (art. 14.2.e del PIDCP). En cumplimiento del artículo 14.2 del PIDCP estas garantías deben aplicarse en igualdad.

Por esa razón, las personas acusadas de delitos deben conocer en una forma que comprendan, el funcionamiento de los sistemas de inteligencia artificial, incluyendo la información de la que se alimentan las bases de datos que permiten automatizar las decisiones. Deben ser informadas también de la forma en que se adoptó la decisión. Es decir, si fue totalmente adoptada por la herramienta de IA o sirvió como un apoyo parcial para orientar la labor judicial. En caso de que la decisión haya sido adoptada enteramente por la IA, las personas acusadas deberían estar en condiciones de presentar evidencia que

¹ Aleš Završnik en el artículo Sistema penal, sistemas de inteligencia artificial y derechos humanos analiza el impacto en el debido proceso de los sistemas de inteligencia artificial desde la perspectiva del Convenio Europeo de Derechos Humanos.

² Carmen Eloísa Ruiz López, et. al. Sistemas Operados mediante Inteligencia Artificial (IA) y debido proceso penal. Perspectiva de aplicación en Colombia, Op. cit., pág. 844.

³ Comité de Derechos Humanos. Caso Sergey Khmelevsky v. Belarús. CCPR/C/139/D/2792/2016. 2024, párr. 8.2.

⁴ Jordi Nieva Fenoll. Inteligencia artificial y proceso judicial, 2019, pág. 153.

cuestione la decisión.

Los problemas para el debido proceso que generan el uso de sistemas basados en IA son similares a los que se presentan cuando se utilizan testigos protegidos o anónimos en procesos penales¹. Esto se afirma debido a la manera en que se programan los sistemas es opaca², no se conoce su financiamiento y quién los diseña³. No conocer el funcionamiento de los algoritmos sería equivalente a obligar a una persona a ser juzgada en un proceso respecto del que conoce las leyes aplicables o en un idioma que desconoce⁴.

Cuando no se garantizan todas estas condiciones, se vulnera la igualdad de medios procesales en perjuicio de la persona acusada⁵ así como el derecho defensa. Lo anterior es así porque debido al desconocimiento del funcionamiento de los sistemas basados en IA⁶ la abogada o el abogado se ve impedido de preparar una defensa técnica.

e. Derecho a recurrir el fallo

Los instrumentos internacionales establecen el derecho de las personas inculpadas al recurrir el fallo. El artículo 14.5 del PIDCP y el artículo 8.2.h) de la CADH consagran este derecho. De acuerdo con la Observación General No. 32 del Comité de Derechos Humanos, el derecho a recurrir el fallo abarca la revisión íntegra de la sentencia condenatoria y de la pena. Dicha revisión va más allá de los aspectos formales o legales de la sentencia. Debe incluir un nuevo análisis de las cuestiones fácticas, probatorias y jurídicas en que se basa la sentencia impugnada⁷. También debe ser realizada por un juez tribunal distinto y de superior jerarquía orgánica⁸. El objetivo de este derecho es evitar que quede firme una decisión adoptada en un procedimiento viciado y con errores que ocasionen un perjuicio al acusado⁹.

Para respetar este derecho, debe asegurarse que las personas acusadas de delitos puedan recurrir íntegramente las decisiones que les afectan y que fueron adoptadas con inteligencia artificial. Es necesario que la revisión sea capaz de analizar las cuestiones fácticas y probatorias en las que se basó la decisión con el fin de identificar posibles errores. Por ejemplo, si una herramienta basada en IA decide que una persona representa un nivel de riesgo alto y con base en ello, se ordena que permanezca en prisión preventiva durante el proceso, la persona debería poder impugnar la decisión. Para ello, tiene derecho a conocer, al menos, qué tecnología fue utilizada, qué información alimentó la base de datos, qué cálculos se hicieron, cómo fue entrenado el algoritmo y qué porcentaje de la decisión correspondió al sistema y qué porcentaje al juez. Estos datos deberían estar disponibles para que una autoridad distinta y de mayor jerarquía revise la decisión.

Si no se asegura que el derecho a recurrir el fallo pueda operar del mismo modo que cuando no se utiliza inteligencia artificial para adoptar una decisión en un proceso penal, no debería permitirse el uso de dicha tecnología¹⁰.

¹ Aleš Završnik, *Criminal justice, artificial intelligence systems and human rights*, 2020, pág. 577.

² Carmen Eloísa Ruiz López, et. al. *Sistemas Operados mediante Inteligencia Artificial (IA) y debido proceso penal. Perspectiva de aplicación en Colombia*, Op. cit., pág. 846.

³ *Ibid.*, pág. 848.

⁴ Jordi Nieva Fenoll, *Inteligencia artificial y proceso judicial*, 2019, pág. 128.

⁵ Comité de Derechos Humanos. *Caso Amir Abdiev v. Kazajstán*. CCPR/C/137/D/2618/2015. 2023, párr. 7.8

⁶ Carmen Eloísa Ruiz López, et. al. *Sistemas Operados mediante Inteligencia Artificial (IA) y debido proceso penal. Perspectiva de aplicación en Colombia*, Op. cit., pág. 849.

⁷ *Caso Mohamed vs. Argentina*

⁸ *Caso Herrera Ulloa vs. Costa Rica*.

⁹ *Caso Barreto Leiva vs. Venezuela*

¹⁰ Jordi Nieva predijo que el uso de la IA en el proceso reconfigurará las impugnaciones o incluso las menguará. Ver. *Inteligencia artificial y proceso judicial*, Op. cit., pág. 144.

f. Independencia e imparcialidad

La independencia y la imparcialidad también deben ser garantizadas cuando se utilizan herramientas basadas en IA en los procesos penales. Estas garantías se analizan tradicionalmente a la luz de la actuación de las personas juzgadoras, pero cuando un sistema basado en AI participa en una decisión en el proceso penal que determina derechos, debe asegurarse la independencia y la imparcialidad desde otras coordenadas, es decir, a partir de las máquinas¹.

De acuerdo con los principios de Bangalore, la independencia judicial requiere que el juez esté libre de conexiones y de influencias inapropiadas con los poderes ejecutivo y legislativo². Estas conexiones e influencias externas tampoco son bienvenidas tratándose de actores privados como empresas.

Asimismo, siguiendo la Observación General No. 32 del Comité de Derechos Humanos, la imparcialidad se refiere a la falta de prejuicio o sesgo del tribunal para tomar una decisión adecuada en un caso. Los principios de Bangalore agregan que la imparcialidad se refiere a la decisión judicial en sí misma y al proceso de la toma de decisión.

Aplicando dichos criterios a las decisiones judiciales adoptadas con inteligencia artificial, debe analizarse, al menos, de dónde proviene la información que alimenta el algoritmo, quién lo diseña y la fuente de financiamiento del sistema. Por ejemplo, podría comprometer la independencia judicial un sistema basado en IA que automatice decisiones judiciales a partir de datos proporcionados por la fiscalía o por la policía. Lo mismo sucedería si una herramienta de AI utilizada por el poder judicial es financiada por la fiscalía o por un actor externo. Además, si los datos automatizados por un sistema de IA reproducen sesgos introducidos por la fiscalía y sirven para alcanzar una decisión judicial, la imparcialidad podría verse comprometida. Es decir, se debe evitar que el poder ejecutivo influya en las decisiones judiciales en donde la IA tiene intervención, así como se hace cuando la IA no interviene³.

Lo anterior se agrava con el hecho de que la incorporación de la IA a la decisión judicial puede encubrirla bajo el velo de la tecnicidad⁴.

8. Conclusiones

Varios gobiernos de Europa están incorporando tecnología basada en inteligencia artificial para realizar labores de seguridad pública y para el funcionamiento del sistema penal. Sin embargo, las investigaciones realizadas demuestran que la inteligencia artificial está implicando un alto costo para los derechos humanos. Por ejemplo, las herramientas de predicción de delitos profundizan la discriminación contra poblaciones en situación de vulnerabilidad o por motivos raciales o por creencias religiosas.

Ese ha sido el caso de la herramienta RADAR-iTe en Alemania y NDAS en el Reino Unido

Además, su efectividad está en entredicho como lo demuestran los errores cometidos por HART en el Reino Unido y ProKid en Holanda.

¹ Jordi Nieva Fenoll. Inteligencia artificial y proceso judicial, Op. cit., pág. 128.

² El Comité de Derechos Humanos señaló que la independencia judicial requiere que el juez y el fiscal puedan "interpretar y aplicar la ley, y evaluar los hechos y pruebas libremente sin ser objeto de intimidaciones, obstrucciones o interferencias en el ejercicio de su función". Comité de Derechos Humanos. Caso Baltasar Garzón v. España. CCPR/C/132/D/2844/2016. 2021, párr. 5.5.

³ Jordi Nieva Fenoll. Inteligencia artificial y proceso judicial, Op. cit., pág. 130.

⁴ Carmen Eloísa Ruiz López, et. al. Sistemas Operados mediante Inteligencia Artificial (IA) y debido proceso penal. Perspectiva de aplicación en Colombia, Op. cit., pág. 854.

En América Latina, la implementación de la inteligencia artificial en las labores de seguridad y en el sistema penal ha sido más paulatina y con reservas. Mientras que las

herramientas de predicción de delitos no son comunes, el reconocimiento facial está proliferando como lo demuestran los casos de Brasil, México, El Salvador y Chile. Ello ha prendido las alertas de diversas organizaciones de derechos humanos especializadas puesto que existen riesgos de que dichas herramientas reproduzcan sesgos y profundicen la discriminación contra poblaciones en situación de vulnerabilidad como las personas afrodescendientes.

En el sistema penal, el uso de la inteligencia artificial ha sido en menor escala. Los casos analizados fueron el Projeto Sócrates en Brasil y Prisma en Colombia. Respecto a Sócrates, que procesa apelaciones, hay preocupaciones de que automatice injusticias sistémicas propias del sistema judicial brasileño. Con relación a Prisma, que evaluaba el riesgo de reincidencia, generó preocupación sus impactos a las garantías del debido proceso como la defensa, la presunción de inocencia y la imparcialidad. Sin embargo, no se tiene información de que se continúe utilizando.

Por otro lado, los ejercicios realizados en Uruguay, Chile y Colombia con sistemas de predicción de delitos dan cuenta de que los gobiernos y las universidades están actuando con precaución en varios países de América Latina. En lugar de implementar la tecnología, han realizado pruebas piloto para identificar errores y posibles sesgos.

Asimismo, los ejemplos de PROMETEA en Argentina y PretorLA en Colombia son buenos ejemplos del uso de la inteligencia artificial para la gestión judicial. En esos países, la automatización está ayudando a hacer frente a problemas habituales de los sistemas de justicia como el retraso en el procesamiento de las causas. Asimismo, en ambos casos se está teniendo conciencia de los riesgos que la inteligencia artificial implican para los derechos humanos.

La inteligencia artificial es una herramienta tecnológica que se está utilizando en los sistemas de justicia de América Latina. No obstante, se debe estar alerta frente al uso de sistemas de decisión automatizada que lejos de lograr los resultados que sus desarrolladores prometen pueden profundizar desigualdades existentes contra las poblaciones más vulnerables.

El poder punitivo del Estado, en tanto implica privar de un bien tan preciado como la libertad a una persona no debe ser sacrificado bajo la promesa falsa de que los sistemas basados en inteligencia artificial son objetivos y toman decisiones libres de sesgos.

Por esa razón, se propone repensar las garantías clásicas del debido proceso como la presunción de inocencia, el derecho de defensa, el derecho a recurrir el fallo y la independencia judicial a la luz de los sistemas de inteligencia artificial. Este informe presenta una primera reflexión que debe profundizarse de tal manera que no se implemente ningún sistema basado en inteligencia artificial que permita violaciones al derecho al debido proceso.

Finalmente, los gobiernos deberían transparentar cualquier plan que incorpore la inteligencia artificial en los sistemas penales. La población general tiene derecho a acceder a esa información en un lenguaje que comprenda. Además, las personas que son sujetas a revisión policial, detenidas, procesadas o condenadas tienen derecho a conocer la forma en que operan los sistemas basados en inteligencia artificial.

9.Recomendaciones

- Desarrollar elementos mínimos para la regulación nacional y regional del uso de la inteligencia artificial. Para ello pueden retomarse buenas prácticas de los países de América Latina así como de Europa como es el caso de la Ley sobre Inteligencia Artificial adoptada por el Parlamento Europeo. Entre ellas:
 - Prohibición el uso de tecnologías de vigilancia masiva como el reconocimiento facial;
 - Prohibición del uso de sistemas de predicción de delitos;
 - Transparencia y rendición de cuentas respecto a la adquisición y uso de sistemas de inteligencia artificial;
 - Desarrollo de garantías del debido proceso aplicables a sistemas de inteligencia artificial durante todas las etapas del proceso penal.
- Recuperar ejercicios como los realizados en Colombia y Chile en los que sin implementar sistemas basados en inteligencia artificial, se realizaron pruebas piloto para evaluar su funcionamiento, los errores y los riesgos que conllevan para los derechos humanos.
- Explorar oportunidades para usar sistemas de inteligencia artificial en las labores de los poderes judiciales con las debidas salvaguardas como la implementación de auditorías externas y con transparencia como en Argentina y en Colombia.
- Realizar auditorías externas a cualquier tecnología que incorpore inteligencia artificial y que sea implementada por cuerpos de seguridad pública o por los sistemas de justicia penal.
- Seguir generando espacios de diálogo e interacción entre autoridades de los sistemas de seguridad pública, poderes judiciales, academia y sociedad civil para discutir los riesgos y oportunidades que presenta la inteligencia artificial para las labores de seguridad pública y del sistema penal.