

Prevención delictiva basada en inteligencia artificial: Reflexiones en torno al derecho a la igualdad

Camila Parodi Ccoriñaupa*

Resumen.- Los sistemas inteligentes para prevenir delitos en el ciberespacio, especialmente agresiones sexuales contra menores enfrentan desafíos entre la eficacia preventiva y la igualdad. La opacidad algorítmica y la ausencia de regulación eficaz facilitan arbitrariedades. Por ello, se debe contar con un marco normativo que armonice adecuadamente la finalidad preventiva y exija la transparencia en su implementación.

Palabras clave.- Igualdad – Ciberespacio – Algoritmos - Discriminación.

Abstract.- Intelligent systems to prevent crimes in cyberspace, particularly sexual offenses against minors, face challenges in balancing preventive effectiveness and equality. Algorithmic opacity and the absence of effective regulation enable arbitrariness. Therefore, a regulatory framework is needed to adequately harmonize the preventive purpose and require transparency in its implementation.

Keywords.- Equality – Cyberspace – Algorithms - Discrimination.

* Estudiante de cuarto año de la carrera de derecho, asistente de cátedra en Derecho Penal y miembro principal del Taller de Dogmática Penal en la Universidad Nacional Mayor de San Marcos.

Nota: En memoria de Joseph Maykol Rolando Rodríguez Chavarry, cuya amistad y cariño seguirán vivos en cada recuerdo.

I. Introducción

El nuevo paradigma de la sociedad quebranta los elementos del tiempo, lugar y espacio conocidos en la narrativa de la humanidad. Este panorama ha derivado del fortalecimiento de las tecnologías de la información y comunicación (TIC), evidenciándose una interdependencia entre la evolución de la sociedad y el desarrollo del conocimiento científico. En este marco, es innegable que estas herramientas ostentan un valor social fundamental en el contexto de la globalización, donde la interacción humana trasciende fronteras y se convierte en un proceso de integración económica, política y cultural. Sin embargo, con el incesante desarrollo, los riesgos ocupan mayor preocupación que las numerosas ventajas de la tecnología.

La diversificación de la criminalidad trae consigo nuevas conductas delictivas que se encuadran dentro del ciberespacio, donde su potencialidad práctica no encuentra límites claros. En este contexto, la delincuencia supera lo conocido por el derecho tradicional y, en consecuencia, exhorta repensar las instituciones jurídicas convencionales.

Todo ello resulta de gran importancia en el mundo de los menores de edad, los sujetos que se encuentran profundamente vinculados con este nuevo escenario. Los nativos digitales desarrollan su personalidad a través de la red, en esta se comunican, informan y entretienen con mayor rapidez que la habitual realidad fáctica. Debido a dicha familiaridad tecnológica, este grupo demográfico se encuentra más propenso a ser víctima de aquellos agresores sexuales que esconden su rostro tras la pantalla, donde gozan de inmediatez y sutileza para perjudicar el adecuado desarrollo de la sexualidad de los menores de edad.

Sumado a ello, una encuesta realizada en el 2023 por el Instituto de Estudios Peruanos (IEP) reveló que más de 500.000 menores de edad recibieron solicitudes para que envíen fotos de su cuerpo y más de 570.000 niñas, niños y adolescentes conocieron a su agresor. Aunado a esto, el Perú es el segundo país en la región con más niños que han entablado conversaciones con extraños en línea. En particular, de acuerdo con el Primer Encuentro de *Grooming* Latam se indicó que a nivel regional el 41% de los casos implica amigos virtuales que solicitaron una relación más allá de una amistad, mientras que en el 21% de los casos, les solicitaron imágenes o videos de contenido íntimo.

Al comparar estas evidencias no cabe duda de que, la situación crítica en la que se encuentra el Perú exhorta adoptar medidas urgentes para contrarrestar estas cifras. Ante ello, surgen diversos problemas respecto a la forma de evitar o prevenir estos actos que afectan la indemnidad sexual, honor y otros derechos fundamentales. Sirva de ilustración los métodos fundados en la concientización social, en la cual se exponen los peligros que se expresan en el entorno digital y se educa a los padres de familia para tener conocimiento sobre cómo actuar ante estas situaciones.

No obstante, en los últimos años surgió la propuesta de combatir estas prácticas mediante la implementación de sistemas de inteligencia artificial. Para ello, se plantea hacer uso de la misma tecnología para identificar la dinámica seguida por los perfiles del espacio cibernético y perseguir a los potenciales agresores sexuales.

A pesar de lo beneficioso que puede resultar, muchos de estos programas –que en diversas partes del mundo están siendo utilizados como policía predictiva– trabajan en base a algoritmos, los cuales automatizan el proceso preventivo y ejecutan una relación de datos que permite un direccionamiento especial de las fuerzas de seguridad. En atención a ello, estos sistemas pueden orientarse por criterios raciales, socioeconómicos y de otra índole que podrían involucrar una diferenciación injustificada, pues se vincula con categorías sospechosas que transgreden el derecho a la igualdad y no discriminación.

Esta investigación pretende exponer la peligrosidad de las agresiones sexuales en el entorno digital y analizar si el uso de sistemas automatizados para la prevención de estas prácticas delictivas es legítimo en la actualidad, todo esto a la luz del derecho a la igualdad.

II. Desarrollo del estado de la cuestión y conceptos fundamentales

i. El ciberespacio como nueva realidad

La Real Academia Española define que el “ciberespacio” es producto de la conjunción de los medios informáticos, en otros términos, se trata de un espacio inmaterial caracterizado por el uso de las redes sociales, los archivos digitales y otras herramientas que permiten la transmisión de información en la red. En ese sentido, se trata de un nuevo ámbito social (una nueva realidad) que facilita la conectividad y el flujo comunicativo,

fortaleciendo la interacción en sociedad de la mano con la inteligencia artificial, el *big data* y elementos propios de esta dimensión¹.

Las descripciones anteriores manifiestan que la existencia del ser humano ya no se encuentra limitada a la materialidad tradicional, sino que se han abierto innumerables posibilidades para la transmisión de ideas, datos y cualquier tipo de información a nivel universal. No es para menos que en el 2019, la Estrategia Nacional de Ciberseguridad (en adelante, ENC) del Reino de España catalogó al ciberespacio como un “espacio común global caracterizados por su apertura funcional y su dinamismo”.

Sin embargo, los infinitos beneficios que puede ofrecer esta versatilidad digital contraponen la aparición de nuevos riesgos en la convivencia social; de modo que, la regulación normativa se encuentra notoriamente desafiada. Al respecto, la ENC distingue que el ciberespacio se asemeja a un “campo de batalla”, esto es, un escenario de confrontación donde la información y la privacidad son activos de incalculable valor, resultando en un contexto complicado para la protección de los derechos fundamentales.

Entre dichos retos se encuentra el desarrollo de nuevas formas de criminalidad, más sutiles y difícilmente localizables. La denominada cibercriminalidad se constituye como un fenómeno que transforma la comprensión habitual de la actividad delictiva y de los mecanismos para su prevención². Como muestra, se evidencia una especial preocupación por los delitos cibernéticos que afectan a niños, niñas y adolescentes, entre los cuales se encuentra el *child grooming* o acoso cibernético³.

¹ TÁVORA, M. (2022). *Vigilancia e investigación policial en el ciberespacio: aspectos procesales del ciberpatrullaje* [Tesis doctoral, Universidad de Sevilla]. Repositorio institucional de la Universidad de Sevilla. <https://dialnet.unirioja.es/servlet/tesis?codigo=309409>

² ROBLES, M. (2025). El ciberespacio y la ciberseguridad: Consideraciones sobre la necesidad de un modelo jurídico. *Instituto Español de Estudios Estratégicos*, (6), pp 1-18.

³ GONZÁLES, V. (2023). *Regulación del delito sexual virtual “Grooming (online o child)” e implementación de inteligencia artificial como posible estrategia para contrarrestar esta figura delictiva en Colombia*. [Tesis de maestría, Universidad de los Andes]. Repositorio Institucional de la Universidad de los Andes. <https://hdl.handle.net/1992/73421>.

Aunado a este panorama, las agendas gubernamentales se encuentran notoriamente influenciadas; puesto que, los Estados requieren fortalecerse ante las crecientes amenazas y propiciar un entorno digital fundado en la confianza y seguridad⁴.

ii. Ciberdelincuencia y violencia digital

Hasta el momento se ha evidenciado que el ciberespacio refleja un marco social que excede las categorías jurídicas tradicionales, denotando la necesidad de replantear el marco regulatorio para adaptarlo a las nuevas formas de interacción.

La ciberdelincuencia circunscribe tanto a los delitos comunes, cuya ejecución se apoya de medios informáticos, como a los delitos informáticos, cuya finalidad es atacar a una red informática o un dispositivo⁵. No toda conducta realizada empleando un instrumento tecnológico es un delito informático⁶. Uno de los criterios para dicha clasificación radica en la necesidad del medio informático para su ejecución.⁷

En vista de lo señalado, la criminalidad informática yace como una de las amenazas más generalizadas actualmente, se extiende a miles de instituciones, empresas y demás sectores de la sociedad⁸. En particular, las amenazas de carácter sexual son manifestación de ciberviolencia, nueva modalidad de perjuicio a través de las tecnologías de la información⁹.

⁴ MORÓN, I. (2021). Riesgos que genera el ciberespacio para los derechos fundamentales. *Revista penal México*, (18), pp. 185-200.

⁵ PEÑA, D. (2023). Ciberdelitos y criminalidad informática: Rol de la prevención en la expansión de la ciberdelincuencia. *Informática y Derecho: Revista Iberoamericana de Derecho Informático (segunda época)*, (13), pp. 57-72.

⁶ VILLAVICENCIO, F. (2014). Delitos Informáticos. *Ius et Veritas*, (49), pp. 284-304

⁷ MAZUELOS, J. (2007). Modelos de imputación en el Derecho penal informático. *Derecho Penal y Criminología*, *Derecho Penal y Criminología* 28, (85), pp. 37-54.

⁸ TÁVORA, M. (2022). *Vigilancia e investigación policial en el ciberespacio: aspectos procesales del ciberpatrullaje* [Tesis doctoral, Universidad de Sevilla]. Repositorio institucional de la Universidad de Sevilla. <https://dialnet.unirioja.es/servlet/tesis?codigo=309409>

⁹ DOMÍNGUEZ, R. (2023). La ciberviolencia en la actualidad: la pantalla como arma letal contra la integridad personal. Dilemas contemporáneos: Educación, Política y Valores. <https://doi.org/10.46377/dilemas.v11i1.3770>.

La violencia digital, ciberviolencia o violencia a través del uso de las TIC puede definirse como “aquella conducta intencionada con la que se causa un daño o un perjuicio a través de las Tecnologías de la Información y de la Comunicación, principalmente el teléfono móvil e Internet”¹⁰. Esta nueva modalidad es ampliamente visualizada en el abuso de las redes sociales, en donde la privacidad se involucra de manera significativa, especialmente, para generar contextos de vulnerabilidad digital¹¹.

Sumado a ello, la VIOLENCIA DIGITAL tiene particularidades que la transforman en un área de especial atención. Entre ellas, se obstaculiza la identificación del agresor, lo cual impide una sanción efectiva. Asimismo, este ambiente permite la preminencia del maltrato en cualquier espacio temporal y, principalmente, goza de sutileza desde la vista de terceros no involucrados directamente, lo cual impide una persecución eficaz¹².

A. Particularidades de los delitos informáticos contra la indemnidad y libertad sexual

La violencia sexual ejercida a través de medios informáticos encuentra diversas manifestaciones, las cuales no discriminan grupos etarios; no obstante, los menores de edad son la población más vulnerable ante estas conductas.

Por un lado, el sexting es una práctica sexual que implica la difusión de imágenes o vídeos de contenido íntimo a través de un ordenador, siendo pasible de sanción penal en los casos donde no haya consentimiento de la víctima¹³. Una técnica utilizada para ello es el craking, donde el agente roba el contenido mediante el acceso

¹⁰ GARCÍA, D. et al. (2011). Violencia a través de las tecnologías de la información y la comunicación en estudiantes de secundaria. *Anales de Psicología / Annals of Psychology* 27, (1), pp. 221-31.

¹¹ VERDEJO, M. (2015). *Ciberacoso y violencia de género en redes sociales: análisis y herramientas de prevención*. Universidad Internacional de Andalucía.

¹² GARCÍA, D. et al. (2011). Violencia a través de las tecnologías de la información y la comunicación en estudiantes de secundaria. *Anales de Psicología / Annals of Psychology* 27, (1), pp. 221-31.

¹³ PALOP, M. (2019). *Protección Jurídica de Menores Víctimas de Violencia de Género a Través de Internet Vulnerabilidad de la Menor en sus Relaciones de Pareja*. Ciberacoso y Derecho al Olvido.

al dispositivo electrónico de su víctima, con la finalidad de difundir o ceder las imágenes sin su consentimiento.

Si bien el sexting puede realizarse dentro de los márgenes de la licitud, habrá situaciones donde se convierta en el insumo para una sextorsión o pornovenganza¹⁴. Este comportamiento es una forma de amenaza, extorsión y chantaje sexual realizado por internet con la finalidad de doblegar la voluntad de la víctima mediante la amenaza de difundir su imagen o vídeo íntimo.

Otra forma de violencia sexual digital es el cyberflashing o ciber exhibicionismo que consiste en el envío de imágenes obscenas a una persona desconocida por la red, sin que se lo haya solicitado. Igualmente, el flaming es aquella conducta dirigida a provocar al receptor a través de mensajes ofensivos e intimidantes¹⁵.

Uno de los fenómenos de trascendencia internacional es el *child grooming*, el cual abarca las estrategias que una persona ejerce para ganarse la confianza de un menor de edad y aprovecharse de su vulnerabilidad, buscando obtener contenido de índole sexual y, en los peores casos, solicitar un encuentro con la víctima¹⁶. Este conjunto de estrategias malintencionadas incluye multiplicidad de actividades de índole sexual, que puede incluir conversaciones, emisión de imágenes y, en lo más severo, la concretización de relaciones sexuales virtuales a través de la webcam o un encuentro físico.

Por ello, se reconoce como un proceso gradual de alta complejidad, cuya efectividad suele depender del enamoramiento o seducción que consigue el groomer respecto del menor de edad. Así, su desarrollo ha sido delimitado en cinco rasgos esenciales: 1) técnicas

¹⁴ Denominado *porn revenge* en la doctrina norteamericana para referirse a estos casos.

¹⁵ DOMÍNGUEZ, R. (2023). La ciberviolencia en la actualidad: la pantalla como arma letal contra la integridad personal. Dilemas contemporáneos: Educación, Política y Valores. <https://doi.org/10.46377/dilemas.v11i1.3770>.

¹⁶ ESPINOSA, J. (2019). Ciberdelincuencia. Aproximación criminológica de los delitos en la red. *La Razón histórica: revista hispanoamericana de historia de las ideas políticas y sociales*, (44), pp. 153-73.

de manipulación y control; 2) dirigido hacia un sujeto vulnerable; 3) realizado en diversas relaciones inter-personales; 4) generar confianza o normalizar la conducta sexual; y 5) busca facilitar una exposición prohibida¹⁷.

En la preparación de este contexto delictivo, el sujeto activo responde a distintas motivaciones; no obstante, son dos las vías más evidentes. En primer lugar, la desensibilización frente a la pornografía adulta, lo cual provoca la búsqueda de comunicaciones sexuales con menores de edad. En segundo lugar, la concretización de fantasías a través de la pantalla, evadiendo el mundo real¹⁸.

Pero ¿qué convierte a estas conductas en un reto para el derecho penal? En buena cuenta, la inmediatez que brinda la virtualidad para que el abusador contacte niños, niñas y adolescentes es incomparable a sus posibilidades en el mundo real. Este espacio le permite estudiar a su víctima y recaudar toda la información que le resulte útil¹⁹. Asimismo, la desterritorialización facilita que el victimario agrede a quienes encuentre vulnerables en el universo digital, sin mayores limitaciones. Esto se vincula con la facilidad de encarnar diferentes perfiles, puesto que, la virtualidad incrementa la dificultad para distinguir al abusador y facilita su interacción con sus víctimas²⁰.

Todo ello sugiere que, las agresiones sexuales se magnifican en este universo, donde el sujeto activo se sirve de las herramientas digitales para facilitar sus objetivos nocivos y gozar de impunidad.

¹⁷ McAlinden, A. (2013). "Grooming" and the Sexual Abuse of Children: Implications for Sex Offender Assessment, Treatment and Management. *Sexual Offender Treatment* 8, (1).

¹⁸ WEFERS, S. et al. (2021). Understanding and Detering Online Child Grooming: A Qualitative Study. *Sexual Offending: Theory, Research, and Prevention* 19, pp. 1-27. <https://doi.org/10.5964/sotrap.13147>.

¹⁹ MIRÓ, F. (2012). El cibercrimen: fenomenología y criminología de la delincuencia en el ciberespacio. *Ediciones Jurídicas y Sociales*. <https://dialnet.unirioja.es/servlet/libro?codigo=756351>.

²⁰ CRUZ, L. (2021). El "Child Grooming" y regulación del delito sexual virtual contra niños, niñas y adolescentes en Colombia. *Derecho Penal y Criminología*, *Derecho Penal y Criminología* 42, (113), pp. 43-96.

De ahí que, el Convenio de Budapest y el conjunto de recomendaciones internacionales ha exhortado a los países a tipificar las diversas conductas que recaen sobre el abuso de menores de edad²¹ en el marco de las nuevas tecnologías.

En esta línea, el uso de sistemas de inteligencia artificial y el análisis predictivo se erige como una medida necesaria para prevenir estos delitos. A través de la policía predictiva, la implementación de algoritmos permite el procesamiento de grandes volúmenes de datos que detectan patrones de riesgo y a posibles agresores. No obstante, existen ciertos riesgos en su desarrollo, lo cual decantaría en falsos positivos y una transgresión injustificada de la igualdad. Por este motivo, el marco normativo internacional se ha adelantado en su regulación.

iii. Regulación normativa sobre el uso de la Inteligencia artificial

La regulación en materia de inteligencia artificial no se ha desarrollado de manera significativa, sin embargo, en los últimos años se tiene esperanza de una evolución continua en la preocupación por este asunto de tanta relevancia, toda vez que involucra una limitación o transgresión sutil de los derechos humanos.

A. Internacional

En primer lugar, se debe enfatizar en los esfuerzos que la Unión Europea (UE) está realizando para elaborar una ley sobre la conducción de la inteligencia artificial, a través de una clasificación de los riesgos y otras normas específicas contextualizadas a la complejidad digital.

Para ello, en el año precedente la UE ha aprobado el Reglamento sobre Inteligencia Artificial, el cual establece normas para regular el desarrollo y uso de sistemas de IA, exhortando que se aplique un enfoque precautorio centrado en el respeto de los derechos fundamentales. Por lo expuesto, se ha realizado una clasificación de los sistemas por el nivel de riesgo, siendo los de riesgo alto aquellos vinculados con derechos como salud, seguridad y otros conexos.

²¹ *Ídem.*

Asimismo, en este documento se especifica el tratamiento que debe seguirse con respecto a la protección de datos personales, señalando que debe evitarse la vigilancia masiva y un uso arbitrario de la identificación biométrica en áreas públicas. Sumado a ello, se destaca la necesidad de implementar el uso de estas tecnologías con transparencia y garantizar la supervisión humana, de manera que se exige que estos sistemas sean supervisables.

En la región americana, Chile busca regular el proyecto de ley sobre sistemas de inteligencia artificial, robótica y tecnologías conexas, considerando la clasificación realizada por la Unión Europea, toda vez que prevé riesgos inaceptables y sistemas de nivel elevado de riesgo. En este país se puede visualizar una evolución significativa en los marcos regulatorios para el uso de la IA, lo cual se hizo visible desde la emisión primigenia de la Política Nacional de Inteligencia Artificial de 2021.

Entre los productos legislativos dirigidos a esta finalidad se encuentran la Ley Marco sobre Ciberseguridad y el Proyecto de Ley sobre Delitos Informáticos que abordan la regulación de la IA en los sistemas de seguridad y en relación con la protección de datos personales. Lo dicho resalta que, la actualización legislativa en Chile se guía por la necesidad urgente de armonizar el desarrollo tecnológico con los derechos fundamentales. Destacando principios de transparencia, no discriminación y rendición de cuentas en el uso de estas nuevas herramientas, alineándose a lo establecido a nivel internacional por la OCDE y la UNESCO²².

Así también, en la experiencia argentina se ha hecho uso de los estándares internacionales en base a principios éticos recomendados por los organismos internacionales anteriormente citados. Entre las principales acciones del Estado argentino, se puede mencionar la inclusión de la IA en la Agenda Digital Argentina 2020 y la elaboración de los Principios Éticos para la IA del 2021. Estos documentos adscriben la equidad, no discriminación, seguridad

²² REMESEIRO, R. (2025). La propuesta de Regulación de la Inteligencia Artificial en Chile y el Reglamento Europeo de Inteligencia Artificial, ¿un caso de efecto Bruselas? *IDP. Revista de Internet, Derecho y Política*, (42), pp. 1-12. <https://doi.org/10.7238/idp.v0i42.431597>.

y responsabilidad en el manejo de las herramientas inteligentes, enfocándose en la prevención.

La trascendencia de estos avances resalta la exigencia de una regulación especial centrada en los riesgos de discriminación algorítmica, vigilancia masiva, decisionismo automatizado y el quebrantamiento general de la privacidad en aras de otros intereses colectivos. Frente a esto, Argentina incorpora un enfoque que regula la IA por su utilidad para el ser humano, lo cual permite priorizar los elementos fundantes de un Estado de Derecho y fortalecer la garantía de los derechos fundamentales sin desconocer la innovación tecnológica²³.

Otro país que se encuentra en proceso de construcción de un marco legal y ético en la utilización de la IA es México, centrándose en la construcción de políticas públicas y participación de organismos especializados. Uno de los avances más destacados es la creación de una Estrategia Nacional de Inteligencia Artificial en 2018, la cual propuso el desarrollo inclusivo, ético y sostenible de las nuevas tecnologías, sin eludir los principios contenidos en la OCDE y la UNESCO.

Pese a todos los esfuerzos, aún se identifican retos importantes para traducir las iniciativas en un marco regulatorio específico sobre el uso de la IA. Aún persiste la necesidad de garantizar la supervisión humana sobre los sistemas inteligentes, en especial dentro de los contextos relacionados a la justicia, la seguridad pública y otros servicios esenciales²⁴.

B. Nacional

El 5 de julio del 2023 se promulgó la Ley N.º 31814, la cual buscó impulsar el desarrollo económico y social permitiendo el uso de

²³ VERCELLI, A. (2024). Regulaciones e inteligencias artificiales en Argentina. *Inmediaciones de la Comunicación* 19, (1), pp. 52-74. <https://doi.org/10.18861/ic.2024.19.1.3549>.

²⁴ MEDINA, M. y TORRES, T. (2025). Regulación de la Inteligencia Artificial: Desafíos para los Derechos Humanos en México. *RIDE Revista Iberoamericana para la Investigación y el Desarrollo Educativo* 15, (30). <https://doi.org/10.23913/ride.v15i30.2291>.

tecnologías y, en específico, la inteligencia artificial, la cual define como una tecnología emergente que tiene un propósito general: potenciar el bienestar humano.

Entre los principios fundamentales que se desarrollan en este instrumento se verifica un enfoque fundado en la identificación del nivel de riesgo, tal como se realizó en el reglamento emitido por la Unión Europea (UE). Igualmente, mediante la pluralidad en la participación se busca que los actores públicos colaboren con los entes privados para asegurar una visión correcta e inclusiva. Esto último encuentra vinculatoriedad en el ejercicio ético y acorde a los derechos fundamentales, toda vez que se busca preservar la privacidad en su uso.

Pese a ello, su reglamentación no sigue la clasificación originalmente planteada por la UE; puesto que, no se han contemplado otros supuestos de riesgo inaceptable como los siguientes:

- Sistemas de evaluación de riesgos y elaboración de perfiles predictivos de delincuencia.
- Sistemas de IA que creen o amplíen bases de datos de reconocimiento facial.
- Sistemas de IA de identificación biométrica remota en tiempo real en espacios públicos

Estas consideraciones son importantes para el presente trabajo de investigación, ya que los presupuestos ignorados en la reglamentación de la ley N.º 31814 habilitan el uso indistinto de sistemas de inteligencia artificial para prevenir o investigar delitos en el contexto digital.

Esta ley aún requiere mejoras para garantizar los marcos de acción que tienen las personas ante riesgos que pudieran lesionar sus derechos fundamentales entre los cuales están vinculadas al acceso, transparencia de la información y protección de datos personales, simplificación de procesos administrativos y otras sobre seguridad informática y gestión de las tecnologías de la información²⁵. Sin

²⁵ FLOR, A. (s. f.) La vulneración de los derechos fundamentales a través de la inteligencia artificial. *SAPIENTIA & IUSTITIA*, (11), pp. 5-24. <https://doi.org/10.35626/sapientia.11.6.113>.

embargo, constituye un avance notorio en la regulación de la IA dentro de la jurisdicción peruana.

iv. Los sistemas inteligentes creados para la prevención y persecución delictiva

La inteligencia artificial (IA) pretende imitar en la mayor medida posible los procesos cognitivos humanos, de tal manera que su aprendizaje busca asemejarse a la percepción y captación de una sumatoria de experiencias. En este sentido, la IA aplicada a la prevención delictiva se basa en el aprendizaje automático o computacional, lo cual permite una construcción de decisiones rápidas en base a contenido preexistente a su operatividad.

En este acápite expondremos algunos de estos sistemas predictivos. Por un lado, HunchLab se centra en identificar áreas predictivas para la criminalidad, lo cual permite que el sistema policial tenga un mejor margen de acción y ostente precisión en su función preventiva. Igualmente, a través del uso de Strategic Subject List se realiza una clasificación de las personas de acuerdo con su historial delictivo, permitiendo que se les asigne un nivel de peligrosidad.

Por otra parte, mediante la implementación de Clearview se permite la identificación automática de un agente delictivo por el reconocimiento facial realizado en base a fotografías disponibles en la red²⁶. Estos sistemas de inteligencia artificial carecen de una visión enfocada en agresiones de carácter sexual, pero son igualmente relevantes de hacer mención.

En el ámbito específico del child grooming, Child Safe AI dirige su funcionalidad a la recopilación de índices de explotación sexual reflejados en los diálogos o fotos emitidos por los usuarios de internet. En el mismo sentido, Spotlight obtiene información de los sitios web que estén vinculados con el tráfico y abuso infantil, donde identifica

²⁶ RODRÍGUEZ, S., MALDONADO, L. (2024). Reflexión sobre el uso de la inteligencia artificial en la investigación penal en el Ecuador-análisis de su aplicabilidad en el derecho comparado. *Polo del Conocimiento* 9, (2). <https://doi.org/10.23857/pc.v9i2.6584>.

a las posibles víctimas, sus agresores y otras conductas dirigidas a esta finalidad delictuosa, esta resulta en una herramienta similar a la Tecnología de IA del Instituto Interregional de las Naciones Unidas para la investigación, el Crimen y la Justicia²⁷.

A su vez, destaca CATT (Chat Analysis Triage Tool), un algoritmo creado por el Instituto Politécnico de la Universidad de Purdue (Indiana), concebido como una herramienta basada en el procesamiento del lenguaje natural. Su finalidad es contrastar conversaciones en línea entre los menores y sus depredadores sexuales para estimar un nivel de riesgo y calcular la probabilidad de que el interlocutor intente concretar un encuentro físico con la víctima.

Por otra parte, una ONG holandesa creó Sweetie, una niña virtual que interactúa con presuntos pederastas en tiempo real, con el objetivo de recabar la información necesaria para su identificación y eventual enjuiciamiento. Su eficacia fue tal que consiguió atraer a más de mil agresores sexuales por internet²⁸.

Lo que se cuestiona de los sistemas de prevención delictiva es que sus inferencias se sustentan en una base de datos que no está del todo verificada y, en consecuencia, puede dar lugar a respuestas sesgadas. Como se ha expuesto, los sistemas son de diversa índole – análisis de texto, procesamiento del lenguaje, etc. – cada uno con sus propios sesgos. Conviene señalar, estos pueden ser de diferente naturaleza: de presentación, de filtro, de muestreo, histórico, de interacción, entre otros igualmente importantes de estudiar²⁹.

Esto se convierte en un gran problema cuando afecta las decisiones policiales y jurisdiccionales, ya que una predicción sesgada puede orientar un despliegue innecesario de diligencias. Asimismo, si la base de datos contiene información específica de ciertos perfiles (idioma,

27 GONZÁLES, V. (2023). *Regulación del delito sexual virtual "Grooming (online o child)" e implementación de inteligencia artificial como posible estrategia para contrarrestar esta figura delictiva en Colombia*. [Tesis de maestría, Universidad de los Andes]. Repositorio Institucional de la Universidad de los Andes. <https://hdl.handle.net/1992/73421>.

28 Noticia: (4 de noviembre de 2013). Niña virtual "atrapa" a más de mil pedófilos, *BBC News Mundo*. https://www.bbc.com/mundo/ultimas_noticias/2013/11/131104_ultnot_paidofilia_virtual_wbm

29 LLANO, F. et al. (2022). *Inteligencia artificial y Filosofía del derecho*. Laborum,

género, origen geográfico, etc.) el sistema tenderá a reproducir dichas características en su análisis de datos.

En el Perú, la Policía Nacional ha elaborado investigaciones sobre las implicancias éticas y legales en torno a la utilidad de la IA y otras herramientas tecnológicas dentro de las operaciones policiales. Un reciente artículo de la Unidad de Posgrado de la PNP concluye que, la ausencia de marcos normativos sólidos y efectivos en la supervisión de la implementación de IA, drones y sistemas biométricos genera riesgos significativos para los derechos fundamentales, dando lugar a prácticas desproporcionadas y arbitrarias³⁰.

Por estos motivos, en la actualidad se contempla la necesidad de estructurar un conjunto de normas que regulen el uso de la tecnología y la inteligencia artificial, en tanto convergen con derechos fundamentales, pero es inevitable su evolución y apropiación a la cotidianeidad.

v. Ventajas y desventajas del uso de los sistemas inteligentes para la prevención delictiva

La principal ventaja radica en su capacidad para procesar y analizar grandes volúmenes de datos con una rapidez y precisión imposibles para la naturaleza humana. Esto permite identificar patrones espaciotemporales y focalizar las fuerzas policiales en los lugares y momentos de mayor probabilidad delictiva (hot-spot policing)³¹. De este modo, herramientas como el reconocimiento biométrico, procesamiento del lenguaje y otros facilitan la evaluación de un sin número de información dispersa.

Otra ventaja de la IA es la personalización de estrategias preventivas y la mejora en la objetividad de las decisiones judiciales. El empleo de la predicción automatizada para calcular índices de reincidencia o peligrosidad hace posible ajustar las decisiones a la realidad de cada caso. Para lograr dicha efectividad, se proponen ciertos filtros, como es exigir que los datos sean ricos en su diversidad, evitando que la representación de ciertos grupos perjudique la predicción neutral.

30 CHAVARRY, L. et al. (2025). Uso de inteligencia artificial, drones y biometría en las operaciones policiales: una revisión sistemática de sus implicaciones éticas y legales. *Revista EspoGra PNP*, 4 (2). <https://doi.org/10.59956/espograpnpv4n2.4>.

31 LUNARDI, et al. (2025). *Digital Transformation and Governance in the Judiciary*. Instituto Jurídico da Faculdade de Direito da Universidade de Coimbra

Esto refiere que, la utilidad será mayor cuando los sistemas incorporen adecuadamente las realidades socioeconómicas, culturales y territoriales en su programación.

Entre las desventajas más significativas se encuentra el riesgo de los sesgos algorítmicos, ya que su implementan depende de las preconcepciones que mantenga su programador. Por esta razón, si los datos reflejan patrones discriminatorios, la IA simplemente amplificará su alcance. Igualmente, el uso de tecnologías como el reconocimiento biométrico remoto, la geolocalización masiva o el rastreo plantea severos cuestionamientos por el tratamiento de información personal sin autorización.

Un claro ejemplo se verificó en un caso estadounidense, donde un ciudadano fue condenado a seis años de pena privativa de libertad, cuyo quantum fue determinado por un software que analiza la reincidencia y permite influir en el marco punitivo³². Dicho sistema fue cuestionado por The Disparate Impact of Criminal Justice Risk Assessment en 2016, pues se comprobó que contaba con algoritmos fundados en categorías raciales, específicamente, personas afrodescendientes a quienes identificaba como de “alto riesgo”.

La operación Sweetie fue la creación de un avatar con características de una niña filipina, con el fin de identificar a los agresores de menores en la red. Este sistema funcionaba de manera automática y con una simulación humana destacable, donde la IA actuaba como una niña y respondía los mensajes adultos que le llegaban a las distintas plataformas en las que interactuaba. Toda la información recolectada era inmediatamente enviada a las autoridades competentes para iniciar investigaciones.

Pese a lo beneficioso que pueda imaginarse, el proyecto generó ciertas discusiones teóricas y éticas, debido a que la recolección de información se realizaba indiscriminadamente, es decir, los datos personales de todos los posibles “sospechosos” estaban en la nube y eran manejados por los organizadores del proyecto, lo cual a primera vista transgrede la privacidad y el uso de datos personales.

En el marco de la actividad policial, se pueden mencionar otros problemas

³² VESTRI, G. (2021). La inteligencia artificial ante el desafío de la transparencia algorítmica: Una aproximación desde la perspectiva jurídico-administrativa. *Revista Aragonesa de Administración Pública*, (56), pp. 368-98.

en los sistemas de IA como la falta de transparencia para la toma de decisiones en la investigación, la opacidad de los algoritmos y su escasa supervisión, lo cual puede dar lugar a trabajar sobre desigualdades preexistentes³³.

La ausencia de regulación sobre estos sistemas se traducirá en decisiones injustas o investigaciones arbitrarias. Por este motivo, los algoritmos y la automatización en la toma de decisiones es un tema de importante atención, toda vez que, los resultados que generan pueden repercutir irremediabilmente en los derechos fundamentales de las personas, no solo el derecho a la igualdad en su dimensión de no discriminación, incluso el derecho a la misma libertad si se empezara a validar su uso y confianza con total plenitud³⁴.

A. Impacto de los sesgos algorítmicos en el derecho a la igualdad

a. Contenido del derecho a la igualdad

La Real Academia Española refiere que el término igualdad alude al principio que reconoce a todos los ciudadanos capacidad para los mismos derechos. No es una concepción alejada de la realidad, no obstante, implica mucho más que considerar a todos como titulares de protección.

En la interpretación constitucional sistemática, algunos derechos presentan una triple identidad: son simultáneamente derechos fundamentales, principios y valores superiores del ordenamiento jurídico. Así, los valores son el fundamento axiológico vinculado a la finalidad última del sistema, los principios son mandatos dirigidos al poder público e incluso a particulares para garantizar ciertos deberes, y los derechos constituyen posiciones subjetivas de ventaja que corresponden a un titular frente a un obligado, implicando la existencia de una prestación exigible.

³³ CHAVARRY, L. et al. (2025). Uso de inteligencia artificial, drones y biometría en las operaciones policiales: una revisión sistemática de sus implicaciones éticas y legales. *Revista Escpogra PNP*, 4 (2). <https://doi.org/10.59956/escpograpnpv4n2.4>.

³⁴ TENORIO, G. (2021). El derecho a una vida libre de algoritmos. *Revista IUS* 15, (48), pp.115-35. <https://doi.org/10.35487/rius.v15i48.2021.708>.

Dicha transversalidad ha sido reconocida para el derecho a la igualdad, reflejando la trascendencia que ostenta dentro del ordenamiento jurídico peruano. En este sentido, la igualdad como principio permite limitar la función legislativa, toda vez que, se constituye como un punto de interpretación³⁵. Esto se condice con el Exp. N.º 0045-2004-PI/TC, en el cual el Tribunal Constitucional ha declarado que este derecho-principio se vincula con la proyección de todo el orden normativo.

En cambio, cuando nos referimos al derecho a la igualdad en su funcionalidad como derecho fundamental implica realizar un análisis de su estructura triádica. Esta particularidad circunscribe tres elementos: una disposición normativa (texto constitucional), norma (interpretación del texto) y una posición iusfundamental (titular-obligado-mandato). Esta última implica la existencia de una relación jurídica, la cual puede plantearse entre individuos o, en otras oportunidades, entre el Estado y los particulares³⁶.

En concordancia, el artículo 2 inciso 2 de la Carta Magna establece que todos los peruanos tienen derecho a la igualdad ante la ley y a no ser discriminado por motivos de raza, religión, opinión socioeconómica u otro motivo análogo. Siendo así, el contenido iusfundamental protegido abarca el derecho a la igualdad en el contenido y aplicación de la ley, derecho a un trato especial cuando se requiera y el derecho a no ser discriminado.

En mayor atención, el derecho a no ser discriminado debe entenderse como una posición iusfundamental de inviolabilidad o inmunidad, toda vez que se configura como una regla inquebrantable. Todo ciudadano tiene el derecho frente al Estado y los particulares de que no se le diferencie por razones

³⁵ LANDA, C. (2021). El derecho fundamental a la igualdad y no discriminación en la jurisprudencia del Tribunal Constitucional del Perú. *Estudios constitucionales* 19, (2), pp. 71-101. <https://doi.org/10.4067/S0718-52002021000200071>.

³⁶ BERNAL, C. (2014). *El principio de proporcionalidad y los derechos fundamentales: El principio de proporcionalidad como criterio para determinar el contenido de los derechos fundamentales vinculantes para el Legislador*. Universidad Externado.

que forman parte de su dignidad como ser humano, a menos que haya una razón constitucionalmente relevante que lo justifique.

De este modo, se proscribe la discriminación recaída sobre “categorías sospechosas”: características inherentes al ser humano (involuntarias) como son la raza, el sexo, idioma, condición económica, orientación sexual, etc. y las posiciones voluntarias que forman parte de la personalidad de un individuo, ya sea el credo, opinión o filiación política.

Cabe precisar que no todo trato desigual es discriminatorio, esto es, si se encuentran motivos objetivos y razonables para la diferenciación, el trato será constitucionalmente tolerable³⁷. El derecho a no ser discriminado coloca a la persona en una posición de tutela especial, donde no debe verse afectado o perjudicado, tratándose de una protección absoluta sobre su esfera individual. Por ende, esta posición iusfundamental no admite una garantía condicionada o sopesada con otros bienes jurídicos³⁸.

- b. Entonces, ¿cómo afectan los algoritmos al contenido del derecho a la igualdad?

Los algoritmos son el conjunto de procedimientos generalizados que se dedican a convertir entradas de datos caóticos en salidas comprensibles mediante una serie de reglas lógicas³⁹. Así parece que la toma de decisiones basada en algoritmos goza de neutralidad y objetividad, pero no es del todo cierto. Para su operatividad se puede valer de desinformación, racismo o prejuicios señalados por su creador⁴⁰.

Si los algoritmos se entrenan con datos sociales sin garantizar su objetividad, pueden generar resultados que reflejen distinciones

³⁷ Sentencia del Tribunal Constitucional 02974-2010-AA, fundamento jurídico 7.

³⁸ SOSA, J. (2024). *El derecho a la igualdad y no discriminación en la jurisprudencia del Tribunal Constitucional y de la Corte IDH*. Palestra.

³⁹ KOED, A. et al. (2016). Big Data: Issues for an International Political Sociology of Data Practices 1. *International Political Sociology*, 10 (3), pp. 275-96. <https://doi.org/10.1093/ips/olw010>.

⁴⁰ LARRONDO, F. (2020). *Inteligencia artificial y derecho*. Hammurabi.

arbitrarias entre distintos grupos de personas. Por ejemplo, el reconocimiento digital que realiza un sistema de inteligencia artificial podría enfocarse en ciertos perfiles que se adecúen a las características tradicionalmente vinculadas con la criminalidad.

Más allá de ser herramientas, los algoritmos constituyen estabilizadores de confianza y garantías fundadas en el mito de la imparcialidad. Sin embargo, no debe ignorarse que se crean con propósitos claros. Determinan preferencias de comportamiento y estructuras para identificar, ordenar y clasificar a las personas⁴¹.

En el marco de la prevención delictiva, uno de los impactos más evidentes de estos sesgos es la discriminación racial y socioeconómica. En la medida que los sistemas de IA están entrenados por personas con sus propias cargas morales incluyen datos históricos que estigmatizan a ciertos grupos. Esto se puede manifestar en aquellos sistemas que agrupan a determinadas etnias como grupos de “alto riesgo” y, en consecuencia, provocan detenciones y cuestionamiento infundados⁴².

Sumado a ello, se ha reconocido el “efecto contagio” delictivo en el uso de estos sistemas de IA, cuyo razonamiento radica en asumir que los delitos futuros serán igualmente ejecutados. Esto significa que, si en un inicio las inferencias brindadas por los algoritmos se dirigieron a señalar determinados perfiles como posibles sospechosos, el mecanismo continuará enfocándose en dichos grupos o en el mismo *modus operandi*. Si no prevé mayores escenarios en la fase de programación, continuará perpetuando la discriminación y eludiendo otras prácticas o personas que se alejen del marco lógico con el que fue creado.

Esto es particularmente relevante en sistemas como CATT, que se enfoca en analizar las interacciones en línea. Si su entrenamiento se ha puntualizado con determinadas pautas culturales, es posible que detecte casos de grooming con mayor precisión en ciertas interacciones y omita otras. Igualmente, con Sweetie, pues al operar como un único perfil digital genera un

⁴¹ KITCHIN, R. (2018). *Thinking critically about and researching algorithms*. Routledge.

⁴² *Ídem*.

sesgo incluso desfavorable para la finalidad preventiva: sesgo selectivo que excluye el perfil de los depredadores sexuales que están en busca de menores de edad con otras características físicas.

Pese a estas falencias, los Estados no pueden quedarse con los brazos cruzados. Si los índices de las agresiones sexuales en línea – o la criminalidad en general – señalan una emergencia, es innegable que se deben implementar medidas lo suficientemente efectivas para contrarrestar estas conductas.

Dado que el derecho a la igualdad y no discriminación posee un contenido de indemnidad frente a diferenciaciones injustificadas, se debe optar por una manera de legitimar el uso de estos sistemas predictivos sin anular el contenido constitucionalmente protegido.

Sobre ello, si se privilegiara únicamente la seguridad pública, se aceptaría sin más la implementación de los sistemas predictivos inteligentes, pese a los evidentes sesgos algorítmicos que presentan. Mientras que, si solo protegemos la igualdad y no discriminación, se impide potenciar la capacidad del Estado para prevenir delitos con alta complejidad, como en estos espacios digitales. Al respecto, uno de los métodos de resolución de conflictos entre derechos fundamentales es el criterio de concordancia práctica, el cual permite la coexistencia de los intereses en conflicto, conservando su máxima eficacia posible⁴³.

A partir de ello, para optimizar el derecho a la igualdad y garantizar la seguridad pública en el espacio digital, se hace necesario introducir una serie de mecanismos que permitan valerse de la utilidad de la inteligencia artificial, sin sacrificar el derecho a la igualdad y no discriminación.

⁴³ RONDON, B. (2024). La concordancia práctica: una relectura a los planteamientos de Konrad Hesse. *Cuestiones constitucionales*. <https://doi.org/10.22201/ijj.24484881e.2024.51.17960>.

B. Medidas para reducir la incidencia de los sesgos algorítmicos

En primer lugar, resaltamos que la naturaleza de los delitos ejecutados en redes informáticas requiere un tratamiento especializado y contextualizado. Con mayor énfasis, la prevención e investigación de las agresiones sexuales en esta nueva realidad es un tema de amplia preocupación.

No obstante, el accionar de los agentes encargados de la prevención e investigación delictiva requiere acogerse a la protección de los derechos constitucionalmente reconocidos. El Estado cuenta con un amplio margen de acción y un poderío que requiere limitaciones precisas y reforzadas, en especial, cuando se trata de implementar medidas automatizadas.

Los ejemplos descritos anteriormente, evidencian el riesgo latente en la operatividad de estos sistemas de inteligencia artificial, su capacidad de predicción se encuentra condicionada a su base de datos, la cual existe sin una regulación especial para identificar el contenido potencialmente lesivo hacia los derechos fundamentales. Por ello, algunos autores los denominan “cajas negras” que no consiguen explicar de qué modo han tomado sus decisiones y eso complica detectar o subsanar los elementos de discriminación que hayan empleado⁴⁴.

No se pretende condenar el uso de sistemas basados en IA, por el contrario, se comprende que se han convertido en una herramienta de amplia utilidad para afrontar la criminalidad. Lo que se busca es exponer la necesidad de una regulación ajustada a su realidad para impedir la arbitrariedad en su uso, atenuar falta de certeza y la comprobada influencia de sesgos algorítmicos.

Como se reconoce en un informe adoptado por la Asamblea Parlamentaria del Consejo de Europa (PACE) en 2017, la peligrosidad de la inteligencia artificial exige regular su fase de programación, con el fin de impedir una injerencia injustificada sobre

⁴⁴ ITURMENDI, J. (2023). La discriminación algorítmica y su impacto en la dignidad de la persona y los derechos humanos: Especial referencia a los inmigrantes. *Revista Deusto de Derechos Humanos*, (12), pp. 257-84.

los derechos humanos y evitar las decisiones fundadas en motivos discriminatorios.

En el contexto peruano, el panorama es aún más preocupante, ya que la citada Ley N.º 31814 no ha previsto las falencias de estos mecanismos inteligentes y abre la posibilidad de su empleo en las diferentes actividades de los órganos públicos y privados. Ante ello, la posible implementación de estos sistemas sin la exigencia de criterios de transparencia y corrección de sesgos algorítmicos permite reforzar las prácticas discriminatorias, lo que es peor, de manera sutil y difícilmente identificable.

Para evitar la falta de certeza y la determinación de los sesgos en la funcionalidad de estos sistemas se deben implementar una serie de estrategias preventivas. Por un lado, diversificar los datos, pues la variabilidad en la data aprendida por los algoritmos permitirá que la información delictiva, por ejemplo, no se encasille en determinadas poblaciones que históricamente se han visto vinculadas a la criminalidad por los prejuicios y la segregación. Paralelamente, debe incluirse un universo de contextos y situaciones posibles para dotarla de mayor efectividad.

Asimismo, el funcionamiento de los algoritmos debe ser accesible para entender los criterios que sujetan las inferencias a las que arribe. De este modo, las auditorías realizadas por terceros independientes sería una buena opción para permitir una evaluación exhaustiva de los datos que se manejen en su programación⁴⁵.

Finalmente, la incorporación de principios como la rendición de cuentas, transparencia u otros construye un marco de acción que paralelamente debe ser reforzado con una campaña de concientización sobre los beneficios y las amenazas en el empleo de la inteligencia artificial.

Estas son algunas de las medidas para mitigar el impacto de los sesgos algorítmicos y hacer uso de la IA con la responsabilidad que se requiere. Valerse de su utilidad y practicidad es oportuno, pero en el marco del respeto hacia los derechos fundamentales.

⁴⁵ BUSTELO, J. (2024). El Sesgo algorítmico en la inteligencia artificial. Implicaciones para la seguridad global y la política criminal. *GEOCONFLICTS & INTELLIGENCE: Revista Sobre Seguridad Global y Terrorismo*, 2 (2), pp.183-203.

III. Conclusiones

El panorama no es tan desalentador como parece, una eventual integración de los sistemas de inteligencia artificial a la vigilancia o prevención delictiva requiere de un marco normativo especializado y estructurado para cumplir con su finalidad y no transgredir las garantías constitucionales.

Los sistemas predictivos pueden verse como una herramienta eficaz en la lucha contra la criminalidad, no obstante, la experiencia comparada ha demostrado que existe una tendencia a reproducir sesgos algorítmicos, los cuales se basan en características sociales injustificadas, “categorías sospechosas”.

El aumento constante de la ciberdelincuencia exige al legislador y los encargados de la persecución delictiva adoptar herramientas tecnológicas que optimicen la detección e investigación de estas nuevas formas de criminalidad. En ese sentido, para brindar legitimidad a estas medidas y no incurrir en la lesión del derecho a la igualdad en la toma de decisiones automatizadas, se debe aplicar el criterio de concordancia práctica y optimizar ambos intereses: la seguridad y el derecho a la igualdad.

Para hacer posible esta armonía, es indispensable implementar exigencias previas como la diversificación de datos en su programación, implementación de auditorías independientes y la concientización en el uso ético de estas herramientas. Si bien no se trata de abandonar las oportunidades de su funcionalidad, la incorporación de sistemas de inteligencia artificial en la prevención e investigación de prácticas criminales como el child grooming requieren cautela y previsión en su regulación.