

TRATA DE PERSONAS EN LA ERA DE LA INTELIGENCIA ARTIFICIAL GENERATIVA: MIRADA DESDE LOS DERECHOS HUMANOS

HUMAN TRAFFICKING IN THE AGE OF GENERATIVE ARTIFICIAL INTELLIGENCE: A HUMAN RIGHTS PERSPECTIVE

Autor: Ana Paula Mendoza Rázuri*

RESUMEN

El presente artículo tiene como objetivo analizar cómo la inteligencia artificial generativa puede facilitar formas de explotación humana cada vez más sofisticadas y a gran escala en entornos digitales, así como examinar los desafíos que esta tecnología emergente plantea para la protección de los derechos humanos. Para ello, la primera sección aborda el papel que han desempeñado históricamente los medios digitales en la facilitación de la trata de personas. La segunda sección examina las capacidades técnicas de la inteligencia artificial generativa y sus posibles usos por parte de redes criminales, destacando su potencial para expandir la escala y el alcance de estas prácticas, incluyendo la exposición de nuevas comunidades previamente fuera del radar. Finalmente, la tercera sección ofrece un estudio exploratorio sobre los incipientes desarrollos en el derecho internacional de los derechos humanos frente a este fenómeno, identificando los desafíos normativos que aún deben abordarse para responder eficazmente a los riesgos que plantea esta tecnología en acelerada evolución.

Palabras clave: Inteligencia artificial generativa - derechos humanos - trata de personas - explotación digital - deepfakes.

ABSTRACT

This article explores how generative artificial intelligence may facilitate increasingly sophisticated and large-scale forms of human exploitation in digital environments, and considers the implications of its misuse for the protection of human rights. The first section examines the historical role of digital technol-

(*) Abogada por la Pontificia Universidad Católica del Perú (PUCP). Máster en Derechos Humanos por la London School of Economics and Political Science (LSE). Investigadora en temas de tecnología y explotación humana en ActiveFence, Londres, Reino Unido. Autora de publicaciones sobre derechos humanos. Contacto: apmendoza@pucp.pe.

ogy in enabling human trafficking. The second analyzes the technological capabilities of generative AI and its potential appropriation by criminal networks, with particular attention to its ability to amplify exploitative practices and reach previously unexposed communities. The third section offers an exploratory analysis of the nascent responses emerging within international human rights law to this evolving phenomenon, highlighting key regulatory challenges that must be addressed to mitigate the risks associated with the rapid development of this technology.

Keywords: Generative artificial intelligence - human rights - human trafficking - digital exploitation - deepfakes.

INTRODUCCIÓN

Históricamente, las redes criminales de trata y explotación de personas han demostrado una gran capacidad de adaptación tecnológica, utilizando estratégicamente los medios digitales para localizar, reclutar, coaccionar y controlar a sus víctimas. Los avances recientes en inteligencia artificial generativa no son la excepción, con claras evidencias de cómo estos actores han incorporado diversas tácticas, como el uso de chatbots multilingües, avatares hiperrealistas o *deepfakes*, para potenciar sus operaciones a gran escala con menores posibilidades de detección.

Este fenómeno cobra especial relevancia ante el acelerado desarrollo de modelos como *Gemini* o *ChatGPT*, capaces de generar imágenes, videos y locuciones sintéticas con un nivel de sofisticación sin precedentes, desdibujando los límites entre lo real y lo artificial. La creciente accesibilidad a estas tecnologías expande las herramientas con las que las redes criminales pueden ejecutar estrategias de explota-

ción en el entorno digital, reconfigurando sus márgenes de acción.

El presente artículo tiene como propósito explorar las formas en las que la inteligencia artificial generativa puede facilitar formas potenciadas de explotación humana, así como examinar los desafíos que esta tecnología emergente plantea para la protección de los derechos humanos desde el derecho internacional.

1. EL ROL DE LOS MEDIOS DIGITALES EN LA FACILITACIÓN DE LA TRATA DE PERSONAS

Desde la invención de Internet y a medida que el entorno digital ha evolucionado, las organizaciones criminales dedicadas a la trata de personas¹ han adaptado progresivamente sus estrategias para aprovechar las oportunidades que brindan las nuevas tecnologías. Según la Oficina de la ONU contra la Droga y el Delito (2019), los medios digitales facilitan todas las fases de la trata de personas, incluyendo la explotación sexual, laboral y criminal, así como

1) De acuerdo con el artículo 3 del Protocolo de Palermo, la trata de personas comprende 3 elementos constitutivos:

- i) **Acción:** Incluye actividades como la captación, el transporte, el traslado, la acogida o la recepción de personas.
- ii) **Medios:** Implica el uso de amenazas, fuerza, coerción de cualquier tipo, secuestro, fraude, engaño, abuso de poder o de una situación de vulnerabilidad. También contempla la entrega o recepción de pagos o beneficios con el fin de obtener el consentimiento de una persona que tenga control sobre otra.
- iii) **Fines:** Tiene como objetivo la explotación, lo cual abarca, como mínimo, la explotación sexual de otras personas, otras formas de explotación sexual, trabajos o servicios forzados, esclavitud o prácticas similares, servidumbre o la extracción de órganos.

En el caso de los niños (personas menores de 18 años), se considera trata aun sin que se empleen estos medios. (ONU, 2000).

“Históricamente, las redes criminales de trata y explotación de personas han demostrado una gran capacidad de adaptación tecnológica, utilizando estratégicamente los medios digitales para localizar, reclutar, coaccionar y controlar a sus víctimas.”

la mendicidad forzada y otras formas de explotación. Esto incluye identificar y captar a las víctimas, promocionar los bienes y servicios resultantes de la explotación, y lavar el dinero obtenido de estos delitos (UNODC, 2019). No obstante, las formas en las que se utilizan los medios digitales varía en cada fase, especialmente en cuanto a la herramienta empleada:

Reclutamiento

El principal canal de contacto para el reclutamiento de víctimas son las redes sociales, ya que representan espacios propicios para acceder a un gran número de personas y ocultar la identidad de los traficantes mediante el uso de perfiles falsos (New Hope Foundation International, 2024). Estas plataformas brindan acceso a información

personal sensible de sus potenciales víctimas —como la edad, ubicación geográfica, intereses, imágenes y redes de contacto— lo que facilita la identificación de vulnerabilidades específicas (UNODC, 2019). A partir de esta información, los traficantes pueden personalizar sus estrategias de manipulación, aumentando así la efectividad del proceso de captación y control (UNODC, 2019). Según la información disponible, Facebook y WhatsApp son plataformas frecuentemente empleadas para la captación de víctimas, aprovechando el anonimato, la información personal compartida por las y los propios usuarios y los canales de comunicación privada que facilitan el proceso de *grooming* (UNODC, 2021).

Control

En la actualidad, muchos casos de explotación pueden desarrollarse sin interacción física directa entre traficantes y víctimas. Tal es el caso de la sextorsión,¹ que consiste en crear una relación de confianza con una potencial víctima, obtener imágenes íntimas, y amenazar con divulgar dichas imágenes para forzar a la víctima a cumplir exigencias económicas, sexuales o de control (Ray & Henry, 2024). Por ejemplo, BM Boys² es un grupo criminal basado en Nigeria y especializado en esta forma de explotación, operando principalmente en Estados Unidos, a través del uso de perfiles de mujeres atractivas para captar víctimas a través de redes sociales (McQue, 2025). Redes transnacionales similares a los BM Boys replican tácticas similares de manipulación y coacción en distintos contextos geográficos. En septiembre de 2025, Organización Internacional de Policía Criminal (INTER-

2) Es importante precisar que la sextorsión y la trata de personas son fenómenos distintos, aunque relacionados. La sextorsión, por sí sola, no configura trata de personas porque, en su manifestación típica, no implica actos de captación, transporte, traslado, acogida o recepción de la víctima con finalidad de explotación. La relación ilícita se circunscribe al chantaje digital orientado a obtener imágenes, actos de naturaleza sexual o dinero, lo que se traduce en un acto de explotación.

3) “BM” hace referencia al término blackmail, que significa “chantaje” en español.

POL) informó la detención de más de 260 presuntos ciberdelincuentes en 14 países africanos, en el marco de una operación de gran escala contra esquemas de sextorsión y estafas románticas en línea (Muia & Booty, 2025). Las principales víctimas de este tipo de redes son adolescentes hombres, y las estadísticas demuestran que este delito puede desencadenar consecuencias letales. Desde el 2021 hasta la fecha, alrededor de 46 víctimas con este perfil en Estados Unidos, Australia, Canadá, y Reino Unido pusieron fin a su vida tras ser víctimas de sextorsión (McQue, 2025).

Explotación

Los sitios web y foros *online* funcionan como plataformas efectivas para la publicidad de víctimas de explotación sexual comercial (New Hope Foundation International, 2024). Según la UNODC (2021), ello abarca tanto páginas especializadas donde trabajadoras sexuales ofrecen sus servicios como sitios de anuncios clasificados. Un caso paradigmático es Backpage.com, un portal de anuncios clasificados que fue incautado por el Buró Federal de Investigaciones de EEUU (FBI) en 2018 debido a su uso sistemático por parte de traficantes para publicitar a víctimas menores de edad (U.S. Department of Justice, 2018). Reportes demuestran que, en la actualidad, existen numerosas plataformas similares en las que se identifican conductas análogas (Bozovic, Novak, 2025).

2. CUANDO LA INTELIGENCIA ARTIFICIAL POTENCIA LA TRATA DE PERSONAS: RIESGOS Y RETOS EMERGENTES

Las herramientas proporcionadas por la inteligencia artificial generativa aumentan significativamente la sofisticación en las tácticas utilizadas por estas redes criminales. Este tipo de tecnología se centra en

“Las herramientas proporcionadas por la inteligencia artificial generativa aumentan significativamente la sofisticación en las tácticas utilizadas por estas redes criminales.”

la creación de contenido nuevo y original, a partir de datos de entrenamiento existentes (Ronge, Maier, & Rathgeber, 2025). Para ello, utiliza modelos de aprendizaje profundo, como las redes neuronales, para aprender patrones y estructuras en grandes conjuntos de datos y generar resultados coherentes y creativos en respuesta a entradas específicas (Noriega, 2024). A continuación, el siguiente cuadro muestra cómo la inteligencia artificial generativa puede producir contenidos multimodales, que comprenden los siguientes contenidos (Fui-Hoon Nah et al., 2023) (ver cuadro en la siguiente página).

Si bien las capacidades de la inteligencia artificial generativa reflejadas en el cuadro líneas arriba vienen transformando positivamente múltiples sectores a través de la automatización y la creación de contenido, estas mismas capacidades también están dando lugar a nuevas formas de uso malicioso. Por ejemplo, investigaciones demuestran que modelos de generación de imágenes como Stable Diffusion han sido utilizados para la creación de imágenes de abuso sexual infantil “hiperrealistas” o como herramienta de “nudificación”, incluyendo escenas crueles de violación de infantes (BBC, 2023). El 2024, el FBI acusó a un

Capacidad generativa	Aplicaciones	Modelos
Inputs/Outputs de Texto	<ul style="list-style-type: none"> • Generación de respuestas a preguntas simples o complejas. • Redacción de artículos, ensayos y otros textos extensos. • Traducción automática entre múltiples idiomas. • Síntesis y resumen de grandes volúmenes de información. • Escritura creativa en diferentes estilos y tonos (cuentos, diálogos, etc.). • Asistencia en la corrección y mejora de textos. 	<ul style="list-style-type: none"> • ChatGPT (OpenAI). • Gemini (Google). • Claude (Anthropic).
Generación de Imágenes	<ul style="list-style-type: none"> • Creación de ilustraciones a partir de descripciones en texto. • Generación de imágenes realistas, artísticas o conceptuales. • Interpretación visual de ideas abstractas o escenas detalladas. • Edición y modificación de imágenes generadas o existentes. 	<ul style="list-style-type: none"> • DALL-E (OpenAI). • Stable Diffusion (Deep Learning). • Veo 3 (Google). • Bing Image Creator (Microsoft).
Generación de Voz	<ul style="list-style-type: none"> • Conversión de texto en voz con entonación natural y expresiva. • Generación de locuciones en múltiples idiomas y acentos. • Lectura automatizada de contenido escrito para accesibilidad o consumo auditivo. • Aplicación en asistentes virtuales, atención al cliente y narraciones. • Voces personalizables, incluyendo clonación de voz o síntesis con emociones. 	<ul style="list-style-type: none"> • ElevenLabs (ElevenLabs) • Amazon Polly (Amazon). • Azure TTS (Microsoft).
Generación de Video	<ul style="list-style-type: none"> • Generación de secuencias de video a partir de descripciones escritas. • Creación de escenas realistas, animadas o estilizadas según el prompt. • Representación visual de conceptos, acciones o entornos complejos. • Aplicación en publicidad, entretenimiento, educación y simulaciones. • Posibilidad de combinar texto, imagen y movimiento en un solo flujo creativo. 	<ul style="list-style-type: none"> • Sora (OpenAI). • Runway Gen-2 (Runway).

hombre de haber producido más de 13,000 imágenes sexualmente explícitas de menores utilizando esta aplicación, en lo que se considera el primer caso en ese país donde las autoridades estadounidenses presentan cargos por la creación de pornografía infantil generada mediante inteligencia artificial (Robins-Early, 2024). La ocurrencia de estos casos es una tendencia en aumento: el Centro Nacional para Menores Desaparecidos y Explotados de EEUU (NCMEC) reportó un incremento del 6,340% de los casos de explotación sexual infantil vinculada al uso de inteligencia artificial generativa entre enero y junio de 2025, con reportes a la CyberTipline que aumentaron de 6,835 a 440,419 (Davis, 2025). Modelos de generación de imágenes similares a Stable Diffusion son susceptibles de ser utilizados con el mismo propósito malicioso, lo que refleja la importancia de implementar filtros y salvaguardas adecuadas.

En el caso específico de la trata de personas, se ha comprobado que las capacidades de reclutamiento y explotación de víctimas mediante el uso de inteligencia artificial generativa trascienden las fronteras geográficas, lingüísticas e incluso los umbrales de veracidad y credibilidad, permitiendo a las redes criminales acceder a un espectro cada vez más amplio de potenciales víctimas a nivel global. Hoy por hoy, esta tecnología permite que cada fase del proceso de captación adquiera una apariencia de legitimidad significativamente superior a la del pasado, incluyendo las interacciones en el idioma nativo de la víctima, la generación de contratos laborales, o la utilización de lenguaje técnico asociado a ciertas profesiones, lo que disminuye la capacidad de las potenciales víctimas para detectar el engaño (Laje & Taylor, 2023). Incluso, los chatbots pueden imitar interacciones humanas con inteligencia emocional, generando confianza en las víctimas y adaptándose a sus señales

afectivas, lo que reduce el esfuerzo manual y continuo que antes requería el proceso de seducción o manipulación (*grooming*) (Makridis, 2025, p. 2).

Dadas las capacidades potenciadas de las redes de trata, la Organización para la Seguridad y la Cooperación en Europa (OSCE) (2024) ha elaborado un marco conceptual que aborda cómo estas organizaciones criminales podrían utilizar algunas de las capacidades clave de la inteligencia artificial generativa para favorecer distintas fases de sus operaciones:

El cuadro evidencia que las capacidades emergentes de la inteligencia artificial amplifican significativamente el alcance de redes criminales dedicadas a la trata de personas, tanto geográfica como culturalmente (OSCE, 2024). Esta tecnología abre camino a procesos más sofisticados, elevando las operaciones de estas redes a escala automatizada y transnacional. Esto expone a comunidades previamente fuera del radar de estos actores, aumentando su vulnerabilidad.

Por ejemplo, un actor criminal basado en Europa podría utilizar un chatbot impulsado por inteligencia artificial para difundir ofertas laborales falsas adaptadas culturalmente, como una en suajili dirigida a trabajadores jóvenes en Kenia para un supuesto puesto de atención al cliente en Dubái, o una en tagalo promocionando un supuesto puesto de enfermería en Canadá dirigida a mujeres con formación en enfermería en Filipinas. De igual modo, la generación de contenido sintético permite la creación de *deepfakes* utilizados para manipular o extorsionar sexualmente a menores, un fenómeno presente en el 11% de los casos reportados de sextorsión infantil (Thorn, 2024). Asimismo, la inteligencia artificial facilita la logística del delito, mediante la producción de documentos falsos –como pasaportes o visas generados con tecnolo-

Capacidad generativa	Tácticas
Traducción en tiempo real de texto y discurso	<ul style="list-style-type: none"> • Usar chatbots para traducir mensajes con precisión cultural y contextual, lo que les permite crear personas en distintos idiomas. • Suplantar a personas de confianza (como familiares o líderes comunitarios) hablando en el idioma nativo de la víctima. • Automatizar la búsqueda en redes sociales en múltiples idiomas y culturas para contactar víctimas en su idioma nativo con mensajes generados por IA culturalmente adaptados.
Generación de publicaciones, ads y personas	<ul style="list-style-type: none"> • Producir anuncios de trabajo realistas y personalizados, adaptados a distintos idiomas, culturas y sectores profesionales • Generar cientos de perfiles falsos que publican testimonios positivos de forma automatizada sobre una oferta laboral o un reclutador, creando una falsa apariencia de legitimidad y popularidad (Astroturfing) • Crear avatares sexualizados con texto, imágenes y voz generados por IA para promocionar servicios sexuales de víctimas de forma encubierta
Generación de deepfake media ⁴	<ul style="list-style-type: none"> • Usar tecnología deepfake para generar identidades falsas, incluyendo imágenes y videos hiperrealistas de personas, con el fin de actuar como reclutadores profesionales o personas buscando una pareja romántica. • Generar contenido sexual no consentido usando la apariencia de una víctima potencial, con el fin de mantener el control.

Elaboración propia. Basado en OSCE (2024).

gía *deepfake*—, lo que dificulta los controles migratorios y la identificación de víctimas (OSCE, 2024).

Lejos de constituir un escenario hipotético, estos casos evidencian manifestaciones reales de la problemática. Un ejemplo reciente es el de los centros de estafa (*scam centers*) en la frontera entre Myanmar y Tailandia, donde Duncan Okindo, sobreviviente de trabajo forzoso, aportó el primer testimonio que vincula directamente el uso de ChatGPT con operaciones de trata y explotación laboral (Global News, 2025; McPherson, 2025). Según su relato, las personas recluidas eran obligadas a utilizar ChatGPT para elaborar mensajes fraudulentos dirigidos a víctimas potenciales

—principalmente agentes inmobiliarios en Estados Unidos—. El propio Duncan fue captado mediante la promesa falsa de un empleo en atención al cliente en Tailandia, lo que muestra que las víctimas también fueron reclutadas mediante métodos engañosos, potencialmente con apoyo de inteligencia artificial.

Este patrón coincide con recientes reportes de INTERPOL, que documentan un uso creciente de la inteligencia artificial generativa en la creación de anuncios laborales falsos destinados a captar víctimas de trata. Asimismo, se ha identificado su aplicación en la generación de perfiles e imágenes sintéticas empleadas en esquemas de sextorsión y estafas románticas (IN-

4) Un *deepfake* se puede entender como un contenido falso generado mediante redes generativas antagónicas, que junto con autoencoders, permiten crear imágenes, videos y audios manipulados (Gautam et al., 2024).

TERPOL, 2025). Esta creciente sofisticación tecnológica plantea serios desafíos para la detección y persecución de la trata. Según Mai (citado en Laje y Taylor, 2023), las autoridades enfrentan una brecha tecnológica y de coordinación frente a redes criminales que emplean herramientas de inteligencia artificial avanzadas y operan con rapidez y anonimato. El volumen y realismo del contenido generado por estas tecnologías superan la capacidad de los sistemas de monitoreo actuales, dado que dicho contenido puede producirse a gran escala, eludiendo los métodos tradicionales de detección (Makridis, 2025, p. 4).

Estos casos ilustran el profundo impacto que la inteligencia artificial generativa puede tener en el ámbito de los derechos humanos cuando es utilizada con fines de explotación. Su capacidad para producir contenido hiperrealista, automatizado y difícil de rastrear facilita enormemente la manipulación, el engaño y la violencia digital a gran escala. Esta tecnología amplifica y automatiza cada una de las fases de la trata de personas, permitiendo que las redes criminales operen con mayor alcance, eficiencia y anonimato.

A pesar de los riesgos señalados, la inteligencia artificial también posee un enorme potencial para fortalecer la prevención y detección de la trata de personas. Si bien Human Rights Watch advierte que el uso de inteligencia artificial en la lucha contra la trata puede ser peligroso debido a los sesgos y estereotipos integrados en los modelos de lenguaje (Kilbride, 2024), estas limitaciones no implican descartar su aplicación. Con marcos éticos sólidos y supervisión humana (*“human in the loop”*),⁵ la inteligencia artificial también puede

“La inteligencia artificial generativa se ha convertido en un asunto de interés creciente en el ámbito del derecho internacional de los derechos humanos, en tanto su impacto presenta un potencial ambivalente.”

ser usada como una herramienta valiosa para combatir la trata desde su origen y a lo largo de sus fases operativas. Investigaciones recientes demuestran que la inteligencia artificial puede emplearse para detectar redes de reclutamiento ocultas. Un estudio de la Universidad de Pensilvania utilizó aprendizaje automático para analizar millones de anuncios en línea y revelar patrones que conectan ofertas laborales falsas con redes de trata sexual, proporcionando a las autoridades instrumentos predictivos para intervenir antes de que ocurra la explotación (Forensic, 2025; Ramchandani, Bastani, & Wyatt, 2025). Otro ejemplo destacado es el uso de reconocimiento facial para identificar víctimas y tratantes en aeropuertos y fronteras, tecnología que INTERPOL y la Oficina de Aduanas y Protección Fronteriza de Estados Unidos (CBP) han empleado con éxito para detectar personas vinculadas a redes de trata en flujos internacionales de viaje (Aguilar Rivera, 2025). En el ámbito financiero, reportes indican que estas he-

5) El concepto de *“human in the loop”* (HITL) hace referencia a un enfoque de inteligencia artificial en el que la intervención humana se integra en el ciclo de toma de decisiones automatizadas para garantizar control, responsabilidad y supervisión ética del sistema (Stryker, 2024).

Derechos	Instrumentos	Vulneración a través de la trata facilitada por inteligencia artificial
Dignidad humana e integridad personal.	DUDH arts. 1 y 5; PIDCP art. 7; CADH art. 5; CEDH art. 3; Carta Africana art. 5	Los avatares y deepfakes generados mediante inteligencia artificial son empleados como instrumentos de engaño, coacción y exposición sexual no consentida, causando humillación y daño psicosocial a las víctimas.
Vida privada, honra y protección de datos	PIDCP art. 17; CADH art. 11; CEDH art. 8	La recolección automatizada de datos (scraping) y el perfilado algorítmico son utilizados para ejecutar un grooming hiperpersonalizado y obtener o fabricar -mediante deepfakes o nudificación- material íntimo con fines de sextorsión. La reindexación y multiplicación de copias impiden el control de las víctimas sobre sus datos personales e imagen, extendiendo la intromisión ilegítima hacia sus entornos familiar, social y laboral.
Prohibición de la esclavitud, servidumbre y trata; libertad y seguridad personales	PIDCP arts. 8 y 9; Protocolo de Palermo art. 3; CADH arts. 6 y 7; CEDH arts. 4 y 5; Carta Africana arts. 5 y 6	La generación automatizada de ofertas laborales fraudulentas, la producción de documentos y credenciales sintéticas, y la suplantación de identidad son empleadas para habilitar la captación, el traslado y la acogida de víctimas mediante engaño o coacción. Durante la fase de explotación, las amenazas de difusión de contenido sintético y los sistemas de vigilancia automatizada restringen la libertad y seguridad personales de las víctimas, perpetuando situaciones de servidumbre o explotación con control transnacional sobre ellas.

Elaboración propia. Basado en OACNUDH (2025) y Thorn (2025).

ramientas vienen siendo utilizadas para rastrear transacciones, detectar patrones anómalos y vincular cuentas cuyos ingresos estarían vinculados a posibles redes de explotación humana (Bagnall et al., 2025).

En suma, estos avances plantean riesgos y oportunidades en igual medida. Esta realidad exige una vigilancia constante de los desarrollos de la inteligencia artificial y una respuesta articulada tanto desde el sistema universal de derechos humanos como de los sistemas regionales, lo cual será explorado en la siguiente sección.

3. DERECHOS HUMANOS ANTE LA EXPLOTACIÓN FACILITADA POR INTELIGENCIA ARTIFICIAL GENERATIVA

La inteligencia artificial generativa se ha convertido en un asunto de interés cre-

ciente en el ámbito del derecho internacional de los derechos humanos (DIDH), en tanto su impacto presenta un potencial ambivalente: por un lado, coadyuva a la promoción y protección de derechos, y por otro, puede ser utilizada para su vulneración a gran escala. Cuando se usa para facilitar la trata de personas, afecta los mismos derechos ya comprometidos por la trata, pero introduce rasgos específicos derivados de sus capacidades técnicas (automatización, alcance ampliado, y replicabilidad del daño). A continuación, se analizan derechos en los que dichos rasgos son más apreciables, sin constituir un listado cerrado (ver cuadro).

A su vez, en el caso de las víctimas menores de edad, se encuentran involucrados los derechos de la niñez. Principalmente relevantes son la protección frente a la explotación y el abuso sexual mediante

grooming automatizado, generación y difusión de deepfakes y procesos de “nudificación” de menores (Convención sobre los Derechos del Niño arts. 19, 32–36).

Bajo estas consideraciones, a continuación, analizaremos, en primer lugar, los avances en el derecho internacional de los derechos humanos vinculados a la trata de personas facilitada por la inteligencia artificial generativa; y, en segundo lugar, algunos aspectos a tomar en cuenta para garantizar el goce de los derechos amenazados por este ilícito:

3.1. Avances en el DIDH

Un balance comparado de los avances sobre la relación entre inteligencia artificial y trata de personas en los distintos sistemas de protección de derechos humanos muestra que el desarrollo normativo y de implementación es aún incipiente. En el sistema universal, los avances han sido principalmente de carácter institucional, orientados a la incorporación del tema en la agenda internacional. Uno de los pasos más relevantes es la creación del Órgano Asesor de Alto Nivel sobre Inteligencia Artificial, constituido en 2023 para realizar tanto el análisis como la formulación de recomendaciones vinculadas con la gobernanza internacional de la inteligencia artificial y su alineación con los derechos humanos y los Objetivos de Desarrollo Sostenible (ONU, 2025). Este órgano ha subrayado el carácter transformador de la inteligencia artificial generativa y ha advertido sobre su potencial de sustitución masiva de mano de obra y los riesgos conexos para el pleno goce de los derechos humanos (ONU, 2025).

Similarmente, otros órganos han iniciado un abordaje preliminar de las implicaciones negativas de la inteligencia artificial especialmente en perjuicio de mujeres y niños/as, como es el caso del material

de abuso sexual infantil y las representaciones pornográficas no consentidas (ONU, 2025). A inicios de 2024, la Relatora Especial de la ONU sobre la venta y explotación sexual de niños advirtió que la inteligencia artificial generativa está facilitando nuevas formas de explotación sexual infantil en línea, incluyendo la creación de imágenes sintéticas como *deepfakes* y *deepnudes* (OACNUDH, 2024). Aquella preocupación también se viene materializando en el plano del Derecho Comparado. En la actualidad, varios Estados—entre ellos, el Reino Unido (Al-Othman, 2024), Estados Unidos (Hughes, 2024), España (Jones, 2024)— han acusado o condenado a individuos por la producción o distribución de material de abuso sexual infantil generado mediante inteligencia artificial, lo que evidencia un contexto político favorable para establecer lineamientos comunes a través del Derecho Internacional respecto al tratamiento de estos ilícitos.

No obstante, el único pronunciamiento que aborda de manera explícita la relación entre la inteligencia artificial y la trata de personas proviene de la Oficina del Alto Comisionado de las Naciones Unidas para los Derechos Humanos (OACNUDH). Recientemente, este órgano publicó la “Taxonomía de los riesgos para los derechos humanos relacionados con la IA generativa”, donde subraya la necesidad de reinterpretar ciertos derechos humanos frente a tecnologías disruptivas como la inteligencia artificial generativa, para garantizar una protección adecuada frente a nuevos riesgos (OACNUDH, 2025). Además, destaca que los impactos negativos de estas tecnologías sobre los derechos humanos tienden a agravarse en poblaciones ya expuestas a situaciones de vulnerabilidad, siendo especialmente preocupante el caso de mujeres y niñas. Dentro de los derechos humanos en riesgo, el derecho en el que la OACNUDH enmarca

los impactos de la inteligencia artificial es el del daño físico y psicológico.⁶

En contraste, en los sistemas regionales de derechos humanos –el Sistema Interamericano (SIDH), el Sistema Europeo (SEDH) y el Sistema Africano de Derechos Humanos y de los Pueblos (SADHP)– no se registran, hasta la fecha, pronunciamientos específicamente orientados a examinar el papel de la inteligencia artificial generativa como medio de facilitación de la trata de personas. Este estudio permite observar que el DIDH aún presenta vacíos en la respuesta integral a los riesgos asociados al uso de la inteligencia artificial generativa, especialmente en lo relativo a su posible uso para facilitar delitos como la trata de personas.

Si bien las medidas existentes en el sistema universal son valiosas, no avanzan al mismo ritmo que el vertiginoso desarrollo de estas tecnologías. La falta de medidas adecuadas y oportunas aumenta la brecha operativa entre las autoridades y las redes criminales, que aprovechan estas tecnologías para escalar la captación y la explotación de sus víctimas a una velocidad que desborda la capacidad estatal de detección, obtención de prueba y cooperación transfronteriza. Este rezago compromete la debida diligencia reforzada que el DIDH impone a los Estados y debilita la tutela de derechos como la dignidad, la integridad, la privacidad y la libertad y seguridad personales, especialmente de mujeres, niñas y otros grupos en situación de vulnerabilidad.

En los sistemas regionales de derechos humanos –Sistema Interamericano (SIDH), Sistema Europeo (SEDH) y Sistema Africano (SADHP)– no se registran, hasta la

fecha, pronunciamientos específicamente orientados a examinar el uso de la inteligencia artificial generativa como medio de facilitación de la trata de personas. De igual forma, no existe jurisprudencia específica, dado que no se han identificado decisiones de tribunales de derechos humanos que aborden directamente esta intersección.

Del análisis realizado se desprende que, hoy por hoy, el DIDH mantiene vacíos para responder de forma integral a los riesgos de la inteligencia artificial generativa, en particular cuando se emplea para cometer o facilitar la trata. Aunque en el sistema universal se han dado pasos valiosos, estos no siguen el ritmo del acelerado desarrollo tecnológico. La ausencia de medidas adecuadas y oportunas agranda la brecha operativa entre autoridades y redes criminales, que aprovechan estas herramientas para escalar la captación y la explotación a una velocidad que desborda la capacidad estatal de detección, obtención de prueba y cooperación transfronteriza. Este rezago compromete el estándar de debida diligencia reforzada que el DIDH exige a los Estados y debilita la tutela de derechos como la dignidad, la integridad, la privacidad y la libertad y seguridad personales, con especial impacto en mujeres, niñas y otros grupos en situación de vulnerabilidad.

3.2. Hacia una respuesta integral

Garantizar una protección efectiva ante riesgos como los que plantea esta tecnología exige una respuesta integral y coordinada que involucre a múltiples actores. En primer lugar, es necesario que tanto el sistema universal como los sistemas regionales de derechos humanos desarrollen

6) Este derecho se encuentra reconocido por diversos instrumentos internacionales: “Toda persona tiene derecho a la vida, a la libertad y a su seguridad” (Declaración Universal de los Derechos Humanos, artículo 3); “Nadie será sometido a torturas ni a penas o tratos crueles, inhumanos o degradantes” (DUDH, art. 5); “Toda persona tendrá derecho a disfrutar del mejor estado posible de salud física y mental” (Pacto Internacional de Derechos Civiles y Políticos, artículo 16); y “Nadie podrá ser arbitrariamente detenido, preso ni desterrado” (PIDCP, art. 9).

estándares específicos que orienten a los Estados frente a los desafíos que plantea esta tecnología en relación con la trata de personas. Estos estándares deben abordar las diversas formas de explotación humana vinculadas al fenómeno, incluyendo no solo la trata con fines de explotación sexual y laboral, sino también manifestaciones emergentes como la sextorsión.

De igual manera, se requiere una actuación proactiva por parte de los Estados para adaptar sus marcos normativos e institucionales relacionados con la trata de personas, incorporando medidas de prevención, protección y sanción frente a las nuevas capacidades de la inteligencia artificial generativa. Esta adaptación debe considerar tanto su dimensión transnacional como su carácter automatizado, y cómo estas tecnologías pueden ser utilizadas con fines delictivos. Estas acciones deben priorizar la identificación de las comunidades que enfrentan mayores riesgos ante formas de explotación potenciadas por el uso indebido de esta tecnología.

Un caso ilustrativo es Estados Unidos, cuyo Departamento de Estado planteó que una respuesta eficaz contra la trata de personas facilitada por la inteligencia artificial exige asumir el carácter de doble uso de esta tecnología: (i) desplegar, en alianza con las plataformas, sistemas de moderación y trazabilidad basados en inteligencia artificial para detectar y bloquear en tiempo real contenidos y anuncios sospechosos (incluidos *deepfakes*), etiquetar su procedencia y remitir alertas para una intervención rápida; y (ii) robustecer las investigaciones con analítica avanzada (p. ej., minería de anuncios de sexo en línea y procesamiento de evidencia digital), junto con capacitación a operadores de justicia bajo un enfoque centrado en la víctima. (U.S. Department of State, 2025).

La participación del sector privado es fundamental para reforzar los esfuerzos contra

la trata de personas facilitada por la inteligencia artificial generativa. Paradójicamente, la tecnología —y en particular la propia inteligencia artificial— representa también una de las herramientas más poderosas para enfrentar esta problemática (U.S. Department of State, 2024). Como sostiene el Grupo de Trabajo sobre empresas y derechos humanos de la ONU, actores de especial atención son las empresas tecnológicas responsables de desarrollar o utilizar los modelos de inteligencia artificial a los que podrían acceder potenciales criminales, para quienes también es aplicable el marco de los Principios Rectores sobre las Empresas y los Derechos Humanos de la ONU, con especial énfasis en el deber de debida diligencia (Grupo de Trabajo sobre empresas y derechos humanos, 2025).

Por ese motivo, es fundamental que estas empresas adopten salvaguardas efectivas en el uso de sus tecnologías, que permitan detectar y prevenir la generación de contenidos susceptibles de ser utilizados para la explotación humana o como medio para facilitar la trata de personas, así como evitar su distribución. Una vez generado el contenido, una medida que puede contribuir significativamente a su trazabilidad es la incorporación de marcas de agua o una etiqueta en materiales creados mediante inteligencia artificial, como imágenes, audio o texto, tal como lo ha implementado Google a través de SynthID (Hulick, K., 2024) o redes sociales como Facebook, Instagram, Threats (Clegg, N., 2024), y TikTok (Nellis, S., 2024). Esta práctica facilita que los usuarios identifiquen cuándo un contenido hiperrealista es de origen sintético, lo cual resulta especialmente útil en contextos donde los *deepfakes* se emplean para difundir ofertas laborales engañosas.

A su vez, las empresas que gestionan redes sociales, reconocidas como canales clave en la captación de víctimas de trata,

no solo tienen la responsabilidad de adoptar medidas preventivas, sino también la capacidad técnica para hacerlo, al contar con modelos propios de inteligencia artificial. Estas tecnologías pueden ser aprovechadas para detectar y frenar dinámicas vinculadas a la trata de personas en sus plataformas. En su Recomendación General N.º 38, el Comité para la Eliminación de la Discriminación contra la Mujer (CEDAW) ha instado a estas compañías a establecer controles eficaces que eviten la exposición de mujeres y niñas a la trata y la explotación sexual a través de sus servicios. De manera específica, el Comité ha llamado a los Estados a “[r]equerir que esas empresas utilicen también su capacidad existente en materia de macrodatos, inteligencia artificial y análisis para detectar cualquier patrón que ayude a descubrir casos de trata y a identificar a las partes involucradas, también en el lado de la demanda” (Comité CEDAW, 2020). En la misma línea, la OSCE ha subrayado que estas plataformas están en condiciones de desarrollar tecnologías de seguridad avanzadas capaces de identificar actividades automatizadas asociadas con la trata, como la creación de perfiles falsos operados por bots que interactúan a través de servicios de mensajería con fines de captación y control (OSCE, 2024).

A la luz de lo expuesto, se incluyen a continuación algunas medidas preliminares que los Estados podrían considerar para la regulación efectiva de la trata de personas y su explotación facilitada por la inteligencia artificial generativa, desde una perspectiva de derechos humanos:

Tipificación agravada de la trata de personas facilitada por inteligencia artificial generativa y formas conexas de explotación

Los marcos normativos penales de los Estados deberán definir las modalidades en las que la trata de personas será conside-

rada como facilitada por inteligencia artificial, incluyendo, por ejemplo, la creación de identidades falsas para captar víctimas, la generación de contenido *deepfake* para su control y la distribución de herramientas diseñadas para la explotación. Asimismo, como forma conexas de explotación, podrá contemplarse la sextorsión. Las penas deben ser disuasorias y armonizar con los sistemas de protección de derechos humanos.

Creación de organismos técnicos especializados

Se recomienda la creación de organismos técnicos especializados encargados de formular políticas integrales para abordar la trata de personas facilitada por inteligencia artificial u otras formas conexas de explotación. Estos organismos deberán contar con recursos técnicos y presupuestarios suficientes para adaptar continuamente las medidas estatales de prevención, investigación y sanción, considerando la rápida evolución de la inteligencia artificial generativa y la fluidez con la que surgen nuevas tendencias en el *modus operandi* de los actores criminales. Asimismo, podrán favorecer la coordinación intersectorial y actuar como observatorios de este fenómeno, generando informes periódicos basados en evidencia sobre su desarrollo a nivel nacional y regional.

En el Derecho Comparado, varios Estados —como España, Argentina, Chile y Brasil (Ulloa, 2024; AESIA, 2025)— han creado órganos técnicos para monitorear el desarrollo de la inteligencia artificial. Aunque no están especializadas en la trata de personas, constituyen un avance institucional beneficioso, a través de la eventual ampliación de competencias para cubrir la prevención y detección de la explotación facilitada por inteligencia artificial, o a través de la creación de unidades específicas de coordinación interinstitucional.

Capacitación de operadores del sistema de justicia

La implementación de tipos penales agravados requiere ser acompañada de programas de capacitación dirigidos a los operadores del sistema de justicia especializados en trata de personas, para que puedan reconocer las manifestaciones específicas de este delito facilitado por inteligencia artificial generativa. Ello fortalecerá su capacidad de identificación, investigación y sanción. También será necesario reforzar sus capacidades técnicas para detectar contenido producido mediante estas tecnologías.

Fiscalización y debida diligencia de las empresas de inteligencia artificial

Se deberá establecer la obligación de que las empresas de inteligencia artificial actúen con debida diligencia, incluyendo requisitos de seguridad tales como la implementación de salvaguardas y filtros específicos contra la generación de contenido que pueda ser utilizado para fines de explotación. Además, deberán presentar reportes sobre los posibles impactos de ciertos algoritmos respecto a niños, niñas y adolescentes, así como protocolos de actuación frente a dichos riesgos.

Programas de concientización dirigidos a la ciudadanía

Será necesario implementar programas de alfabetización digital para informar a la ciudadanía, con especial atención a niños, niñas y adolescentes, sobre los riesgos asociados a la inteligencia artificial generativa, incluyendo su potencial para facilitar la trata de personas, la producción de material de abuso sexual infantil y las representaciones pornográficas no consentidas. Estos programas deberán incluir herramientas para detectar contenidos creados mediante inteligencia artificial y explicar las consecuencias jurídicas y sociales del

uso de estas tecnologías con fines de explotación (U.S. Department of State, 2025).

4. CONCLUSIÓN

La inteligencia artificial generativa presenta riesgos únicos y sin precedentes para los derechos humanos y, en particular, para la lucha contra la trata de personas. Al permitir la generación automatizada de contenido hiperrealista, la creación de identidades falsas y la interacción fluida en múltiples idiomas, esta tecnología amplifica la capacidad de las redes criminales para reclutar, manipular y explotar víctimas en una escala transnacional y automatizada. Estos riesgos se intensifican para las comunidades más vulnerables, especialmente mujeres, niñas y adolescentes, que históricamente han sido las más afectadas por este fenómeno. Pese a su creciente relevancia, la intersección entre inteligencia artificial generativa y trata de personas continúa siendo un campo poco estudiado y con claros vacíos normativos y jurisprudenciales, particularmente en los sistemas de protección de derechos humanos.

Este estudio se plantea como un ejercicio exploratorio cuyo objetivo principal es poner el tema sobre la mesa, visibilizar un campo emergente de riesgos y fomentar un diálogo informado entre sectores clave: derechos humanos, tecnología, y política pública. Los esfuerzos pendientes requieren una acción conjunta y coordinada, en la que la inteligencia artificial, utilizada de forma ética y estratégica, puede convertirse en una aliada para enfrentar estos desafíos. A su vez, requiere una pronta respuesta estatal, para lo cual en este estudio se incluyen recomendaciones preliminares con el objetivo de guiar la formulación de medidas normativas e institucionales capaces de mitigar los riesgos emergentes asociados a estas tecnologías, en consonancia con las obligaciones de derechos humanos frente a la trata de personas. ♦

BIBLIOGRAFÍA

- Agencia Española de Supervisión de la Inteligencia Artificial (AESIA). (2025). *Ensuring ethical and responsible AI*. Financiado por la Unión Europea. <https://aesia.digital.gob.es/en/es>
- Aguilar Rivera, O. R. (2025). Inteligencia artificial para la detección y prevención de la trata de personas. *Revista Mexicana de Ciencias Penales*, 8(26), 3–26. <https://doi.org/10.57042/rmcp.v9i26.895>
- Al-Othman, H. (2024, 28 de octubre). *Man who used AI to create child abuse images jailed for 18 years*. The Guardian. <https://www.theguardian.com/uk-news/2024/oct/28/man-used-ai-create-child-abuse-images-jailed-18-years-hugh-nelson>
- Bagnall, C., Krupena, S., Kemp, C., & Massimi, F. III. (2025, 27 de enero). AI's role in combating human trafficking in the financial sector. ACAMS Today. <https://www.acamstoday.org/ai-role-in-combating-human-trafficking-in-the-financial-sector/>
- BBC News. (2023, 12 de octubre). AI image tool Stable Diffusion used to create child sexual abuse images. <https://www.bbc.co.uk/news/uk-65932372#:~:text=The%20makers%20of%20the%20abuse,in%20art%20or%20graphic%20design>
- Bozovic, Novak. (2025, 24 de abril). *10 Best Backpage Alternatives in 2025*. TechNadu. <https://www.technadu.com/best-backpage-alternatives/91900>
- Celona, L., & Fonrouge, G. (2018, 28 de mayo). Backpage closure hasn't made it any harder for sex traffickers. New York Post. <https://nypost.com/2018/05/28/backpage-closure-hasnt-made-it-any-harder-for-sex-traffickers/>
- Clegg, N. (2024, 6 de febrero). *Labeling AI-Generated Images on Facebook, Instagram and Threads*. Meta Newsroom. <https://about.fb.com/news/2024/02/labeling-ai-generated-images/>
- Comité CEDAW. (2020). Recomendación general núm. 38 (2020), relativa a la trata de mujeres y niñas en el contexto de la migración mundial. CEDAW/C/GC/38.
- Consejo de Europa. (1950). *Convenio para la Protección de los Derechos Humanos y de las Libertades Fundamentales* [CEDH].
- Davis, P. (2025, September 4). Spike in online crimes against children a “wake-up call.” National Center for Missing & Exploited Children (NCMEC). <https://www.missingkids.org/blog/2025/spike-in-online-crimes-against-children-a-wake-up-call>
- Forensic. (2025, 16 de junio). New AI research reveals hidden human trafficking recruitment networks. Compare Networks. <https://www.forensicmag.com/articles/new-ai-research-reveals-hidden-human-trafficking-recruitment-networks>
- Fui-Hoon Nah, F., Zheng, R., Cai, J., Siau, K., & Chen, L. (2023). Generative AI and ChatGPT: Applications, challenges, and AI-human collaboration. *Journal of Information Technology Case and Application Research*, 25(3), 277–304. <https://doi.org/10.1080/15228053.2023.2233814>

- Gautam, A., Joshi, R. K., Narula, A., & Sharma, N. (2024). *Mitigating human rights violations caused by deepfake technology*. *LIB PRO*, 44(3).
<https://doi.org/10.48165/bapas.2024.44.2.1>
- Global News. (2025, 15 de septiembre). *'ChatGPT was helping': Kenyan human trafficking survivor reveals AI role in Myanmar scam centres* [Video]. Global News. <https://globalnews.ca/video/11429097/chatgpt-was-helping-kenyan-human-trafficking-survivor-reveals-ai-role-in-myanmar-scam-centres>
- Grupo de Trabajo sobre empresas y derechos humanos. (2025, 22 de enero). *Convocatoria de aportaciones: El uso de la inteligencia artificial y los Principios Rectores sobre empresas y derechos humanos*. Oficina del Alto Comisionado de las Naciones Unidas para los Derechos Humanos (OACNUDH).
<https://www.ohchr.org/es/calls-for-input/2025/call-inputs-use-artificial-intelligence-and-un-guiding-principles-business-and-human-rights>
- Hulick, K. (2024, 11 de diciembre). *Google now adds watermarks to all its AI-generated content*. *Science News Explores*.
<https://www.snexplores.org/article/google-ai-watermarks>
- Hughes, C. (2024, 1 de mayo). Pittsburgh man gets 14 years for possession of deepfake child abuse images. United Press International.
https://www.upi.com/Top_News/US/2024/05/01/deepfake-child-abuse-sentencing/4451714583861/
- INTERPOL. (2025, 30 de junio). *INTERPOL releases new information on globalization of scam centres*.
<https://www.interpol.int/en/News-and-Events/News/2025/INTERPOL-releases-new-information-on-globalization-of-scam-centres>
- Jones, S. (2024, 9 de julio). *Spain sentences 15 schoolchildren over AI-generated naked images*. The Guardian. <https://www.theguardian.com/world/2024/jul/09/spain-sentences-15-schoolchildren-over-ai-generated-naked-images>
- Kilbride, E. (2024, 1 de julio). *Using AI to fight trafficking is dangerous: Language models built on stereotypes will exacerbate harm*. Human Rights Watch.
<https://www.hrw.org/news/2024/07/01/using-ai-fight-trafficking-dangerous>
- Laje, D., & Taylor, N. (2023, 1 de noviembre). *Technology's good and evil: Fueling and fighting human trafficking*. SIGNAL Magazine. AFCEA International.
<https://www.afcea.org/signal-media/cyber-edge/technologys-good-and-evil-fueling-and-fighting-human-trafficking>
- Makridis, C. A. (2025). *Countering human trafficking risks of generative AI with trustworthy AI and education*. Journal of Human Trafficking. Advance online publication. <https://doi.org/10.1080/23322705.2025.2535986>
- McPherson, P. (2025, 15 de septiembre). *ChatGPT was used "to help scammers do their thing" in Asia fraud scheme*. Reuters.
<https://www.reuters.com/investigations/chatgpt-was-used-help-scammers-do-their-thing-asia-fraud-scheme-2025-09-15/>
- Mufarech, A., & Lung, N. (2025, 3 de marzo). *Sex traffickers in Colombia are using Facebook, Tinder and Airbnb to exploit mi-*

- nors. Bloomberg. <https://www.bloomberg.com/news/features/2025-03-03/facebook-tinder-airbnb-apps-are-used-for-sex-trafficking-in-colombia>
- Muia, W., & Booty, N. (2025, 26 de septiembre). *Mass arrests for sextortion and romance scams in sting operation across Africa*. BBC News. <https://www.bbc.co.uk/news/articles/c0r04vjvkljo>
- Naciones Unidas. (1948). *Declaración Universal de Derechos Humanos* [DUDH]. Oficina del Alto Comisionado de las Naciones Unidas para los Derechos Humanos.
- Naciones Unidas. (1966). *Pacto Internacional de Derechos Civiles y Políticos* [PIDCP]. Oficina del Alto Comisionado de las Naciones Unidas para los Derechos Humanos.
- Naciones Unidas. (1989). *Convención sobre los Derechos del Niño* [CDN]. Oficina del Alto Comisionado de las Naciones Unidas para los Derechos Humanos.
- Naciones Unidas. (2000). *Protocolo para Prevenir, Reprimir y Sancionar la Trata de Personas, Especialmente Mujeres y Niños, que complementa la Convención de las Naciones Unidas contra la Delincuencia Organizada Transnacional* [Protocolo de Palermo]. Oficina del Alto Comisionado de las Naciones Unidas para los Derechos Humanos.
- New Hope Foundation International. (2024, 19 de junio). *The dark side of the web: How technology fuels human trafficking*. <https://nhfinternational.org/how-technology-fuels-human-trafficking/>
- Noriega, H. (2024, 20 de octubre). ¿Qué es IA generativa? Guía básica para entender cómo funciona, beneficios, riesgos y perspectivas. <https://hiramnoriega.com/64226/que-es-ia-generativa-definicion-como-funciona/>
- Oficina del Alto Comisionado de las Naciones Unidas para los Derechos Humanos [OACNUDH]. (2014). *Los Derechos Humanos y la Trata de Personas* (Ficha informativa N.º 36). https://www.ohchr.org/sites/default/files/Documents/Publications/FS36_sp.pdf
- Oficina del Alto Comisionado de las Naciones Unidas para los Derechos Humanos [OACNUDH]. (2024, 5 de febrero). *UN expert alarmed by new emerging exploitative practices of online child sexual abuse*. <https://www.ohchr.org/en/press-releases/2024/02/un-expert-alarmed-new-emerging-exploitative-practices-online-child-sexual>
- Oficina del Alto Comisionado de las Naciones Unidas para los Derechos Humanos [OACNUDH]. (2025). *A taxonomy of human rights harms from generative AI*. <https://www.ohchr.org/sites/default/files/documents/issues/business/b-tech/taxonomy-GenAI-Human-Rights-Harms.pdf>
- ONU. (2000). *Protocolo para Prevenir, Reprimir y Sancionar la Trata de Personas, Especialmente Mujeres y Niños* (Protocolo de Palermo). <https://www.ohchr.org/es/professionalinterest/pages/protocoltraffickinginpersons.aspx>
- ONU. (2024). *Gobernanza de la inteligencia artificial en beneficio de la humanidad: Informe final*. Recuperado el 1 de junio de 2025, de https://www.un.org/sites/un2.un.org/files/governing_ai_for_humanity_final_report_es.pdf

- ONU. (2025). *Inteligencia artificial*. Recuperado el 1 de junio de 2025, de <https://www.un.org/es/global-issues/artificial-intelligence>
- Organización de Estados Americanos. (1969). *Convención Americana sobre Derechos Humanos [CADH]*. Organización de los Estados Americanos.
- OSCE. (2024). *New frontiers: The use of generative artificial intelligence to facilitate trafficking in persons*. Bangkok/Vienna. <https://www.osce.org/files/f/documents/7/d/579715.pdf>
- Phelan, M. (2025, 14 de mayo). Everything Google announced at I/O 2025. WIRED. <https://www.wired.com/story/everything-google-announced-at-io-2025/>
- Ramchandani, P., Bastani, H., & Wyatt, E. (2025). Unmasking human trafficking risk in commercial sex supply chains with machine learning. *Manufacturing & Service Operations Management*, 27(3). <https://doi.org/10.1287/msom.2022.0304>
- Ray, A., & Henry, N. (2024). Sextortion: A Scoping Review. *Trauma, Violence, & Abuse*, 26(1), 138-155. <https://journals.sagepub.com/doi/10.1177/15248380241277271>
- Robins-Early, N. (2024, 21 de mayo). US man used AI to generate 13,000 child sexual abuse pictures, FBI alleges. *The Guardian*. <https://www.theguardian.com/technology/2024/may/21/us-man-used-ai-generate-child-sexual-abuse-pictures-fbi>
- Ronge, R., Maier, M., & Rathgeber, B. (2025). *Towards a definition of generative artificial intelligence*. *Philosophy & Technology*, 38(31), 1–25. <https://doi.org/10.1007/s13347-025-00863-y>
- Thomson Reuters. (2025, 2 de enero). *Technology and human trafficking: Fighting the good fight*. <https://www.thomsonreuters.com/en-us/posts/legal/technology-and-human-trafficking/>
- Thorn. (2024). *Trends in financial sextortion: An investigation of sextortion reports involving minors*. https://info.thorn.org/hubfs/Research/Thorn_TrendsInFinancialSextortion_June2024.pdf
- Thorn. (2025, 3 de marzo). *Deepfake nudes & young people: Navigating a new frontier in technology-facilitated nonconsensual sexual abuse and exploitation*. https://info.thorn.org/hubfs/Research/Thorn_DeepfakeNudes&YoungPeople_Mar2025.pdf
- Oficina de las Naciones Unidas contra la Droga y el Delito [UNODC]. (2019). *Technology that facilitates trafficking in persons*. United Nations Office on Drugs and Crime – Education for Justice. <https://www.unodc.org/e4j/es/tip-and-som/module-14/key-issues/technology-that-facilitates-trafficking-in-persons.htm>
- Oficina de las Naciones Unidas contra la Droga y el Delito [UNODC]. (2021). *Global report on trafficking in persons 2020: Chapter V – The impact of COVID-19 on trafficking in persons*. https://www.unodc.org/documents/data-and-analysis/tip/2021/GLO-TiP_2020_Chapter5.pdf
- Ulloa, M. (2024, 1 de octubre). *Avances en la regulación de la Inteligencia Artifi-*

- cial en América Latina*. Observatorio de Riesgos Catastróficos Globales (ORCG). <https://orcg.info/articulos/avances-en-la-regulacin-de-la-inteligencia-artificial-en-amrica-latina>
- Unión Africana. (1981). *Carta Africana de Derechos Humanos y de los Pueblos* [Carta Africana]. Unión Africana.
- U.S. Department of Justice. (2018, 9 de abril). *Justice Department leads effort to seize Backpage.com, the Internet's leading forum for prostitution ads*. <https://www.justice.gov/archives/opa/pr/justice-department-leads-effort-seize-backpage-com-internet-s-leading-forum-prostitution-ads>
- U.S. Department of Justice. (2024, 20 de mayo). *Man arrested for producing, distributing, and possessing AI-generated images of minors engaged in sexually explicit conduct* [Comunicado de Prensa]. <https://www.justice.gov/opa/pr/man-arrested-producing-distributing-and-possessing-ai-generated-images-minors-engaged-sexually>
- U.S. Department of State. (2024). *2024 Trafficking in persons report*. Office to Monitor and Combat Trafficking in Persons. <https://www.state.gov/reports/2024-trafficking-in-persons-report/>
- U.S. Department of State. (2025). *2025 Trafficking in Persons Report*. Office to Monitor and Combat Trafficking in Persons. <https://www.state.gov/reports/2025-trafficking-in-persons-report/>
- Withnall, A. (2016, 20 de noviembre). *The fight to shut down Backpage.com, America's biggest online brothel*. The Independent. <https://www.independent.co.uk/news/world/americas/the-fight-to-shut-down-backpage-com-americas-biggest-online-brothel-a7411066.html>