EL PACCTO 2.0

EU-LAC Partnership on justice and security

RECLUTADOS POR EL ALGORITMO:

CÓMO LAS REDES CRIMINALES UTILIZAN LA IA PARA RECLUTAR EN LA UE, AMÉRICA LATINA Y EL CARIBE

2026

Financiado por la UE

Coordinado por:

EXPERTISE FRANCE
GROUPE AFD

EMPACT

Con la dirección y revisión de:

Marc Reina Tortosa, Senior Executive Manager de EL PACCTO 2.0 y Colíder de la Acción Operativa 6.1 del EMPACT OAP MTCNI

Sophie Bailly, Comisaria de Policía y Líder de la Acción 1.6 del EMPACT OAP MTCNI

Autor:

Joel Levesque

Con la dirección y revisión de:

Australian National University (Australia)
Bundeskriminalamt (Alemania)
The Caribbean Community (CARICOM)
Implementation Agency for Crime and Security (IMPACS)
Casa Grande University (Ecuador)
El Colegio de México (México)
Europol. Operational Taskforce GRIMM / EMPACT
Police Judiciaire Fédérale (Bélgica)
Policía Nacional (España)
Polisen (Suecia)

Diseño:

Carlos Múgica

DOI: 10.5281/zenodo.20528562

Edición no comercial
Julio de 2026

ÍNDICE

INTRODUCCIÓN

El Laboratorio de Innovación EL PACCTO 2.0 lleva desde 2024 analizando la intersección entre la inteligencia artificial y el crimen organizado, publicando su primer estudio en profundidad sobre el tema y convocando la primera conferencia birregional sobre IA y delincuencia. En 2025, el programa elaboró otros dos informes: «La IA como arma: cómo la IA está transformando el mundo» y «El uso de la inteligencia artificial por parte de redes criminales de alto riesgo». Este estudio da continuidad a ese trabajo, examinando una dimensión que ha recibido menos atención sistemática: el uso de la IA por parte de organizaciones criminales para reclutar a niños, adolescentes y adultos en América Latina, el Caribe y la Unión Europea.

El reclutamiento es la puerta de entrada a prácticamente todas las formas de actividad delictiva organizada.1 La IA lo está transformando de formas que inciden en la aplicación de la ley y las políticas; aumentando la velocidad, la escala y el alcance geográfico de las operaciones de reclutamiento; reduciendo el nivel de competencia necesario para llevarlas a cabo; y permitiendo a los actores delictivos segmentar y dirigirse a poblaciones con una precisión que antes no era posible.

Las herramientas de IA reducen las barreras de reclutamiento a gran escala, pero las capacidades implicadas son más específicas de lo que sugiere la afirmación general sobre velocidad y alcance. Los algoritmos de recomendación de las plataformas amplifican el contenido engañoso dirigido a audiencias específicas. El análisis de datos permite identificar a personas que muestran dificultades económicas, aislamiento social u otros indicadores de

1 En este informe se entiende que el reclutamiento abarca toda la gama de métodos mediante los cuales las organizaciones delictivas incorporan a personas a actividades delictivas, ya sea mediante el engaño, la coacción, incentivos económicos o el consentimiento inducido.

vulnerabilidad. Los mensajes basados en grandes modelos de lenguaje adaptan el tono y el lenguaje al destinatario. El audio y el vídeo sintéticos suplantan a contactos de confianza o inventan ofertas de trabajo convincentes. Los sistemas de seguimiento automatizados mantienen el contacto a través de múltiples plataformas simultáneamente. La línea entre la infraestructura legítima de las plataformas y su explotación delictiva suele ser difusa.

Los ecosistemas delictivos de América Latina y el Caribe (ALC) y la UE presentan diferencias que determinan el funcionamiento del reclutamiento asistido por IA en cada contexto. En la UE, los patrones predominantes consisten en la captación de menores y jóvenes adultos para el tráfico de drogas, el blanqueo de dinero, la «violencia como servicio» y la explotación sexual infantil en línea, principalmente a través de las redes sociales, los entornos de videojuegos y las aplicaciones de mensajería cifrada. En algunas zonas de ALC, especialmente donde la presencia del Estado es débil, el reclutamiento forzoso bajo amenaza física sigue siendo significativo, junto con métodos engañosos en línea. Los cárteles de tráfico de drogas a gran escala, las redes de centros de estafa y las organizaciones afiliadas a grupos paramilitares y guerrilleros presentan ecosistemas delictivos que no se dan en la UE, y sus operaciones de reclutamiento reflejan diferentes grados de coacción y organización.

En ambas regiones, la edad del objetivo determina cómo funciona el reclutamiento, aunque no siempre de forma deliberada. Las organizaciones criminales explotan a quienquiera que sea accesible y útil. La edad se correlaciona con ambos aspectos: determina qué plataformas utiliza una persona, qué argumentos psicológicos son eficaces, a qué consecuencias legales se enfrenta un reclutador si es detenido y qué funciones delictivas puede desempeñar un recluta. El resultado es una segmentación de facto « » observable en los patrones de los casos documentados, incluso cuando no existe una estrategia de selección explícita.

Este informe utiliza la edad como marco analítico más que como una clasificación jurídica fija, dado que las definiciones de minoría de edad y responsabilidad penal varían entre las jurisdicciones de América Latina y el Caribe (ALC) y la UE. En los casos de reclutamiento potenciado por la IA, se distinguen tres grandes grupos. Los niños pequeños se encuentran predominantemente como víctimas de explotación laboral y sexual, donde la dependencia y la incapacidad para dar un consentimiento informado son en sí mismas la vulnerabilidad explotada. Los niños de más edad y los adolescentes son los objetivos documentados con mayor frecuencia en la gama más amplia de tipologías delictivas, y su responsabilidad penal reducida es una ventaja operativa documentada que las redes delictivas explotan deliberadamente. Los adultos son blanco de este fenómeno más comúnmente a través de la vulnerabilidad económica, siendo las ofertas de empleo engañosas, los incentivos económicos y las estructuras de deuda coercitivas los mecanismos principales. La IA potencia el reclutamiento de manera diferente en cada etapa: mediante el grooming, la recomendación de contenidos y el contacto a través de plataformas para los objetivos más jóvenes, y mediante infraestructuras de fraude, identidades sintéticas y manipulación financiera para los adultos.

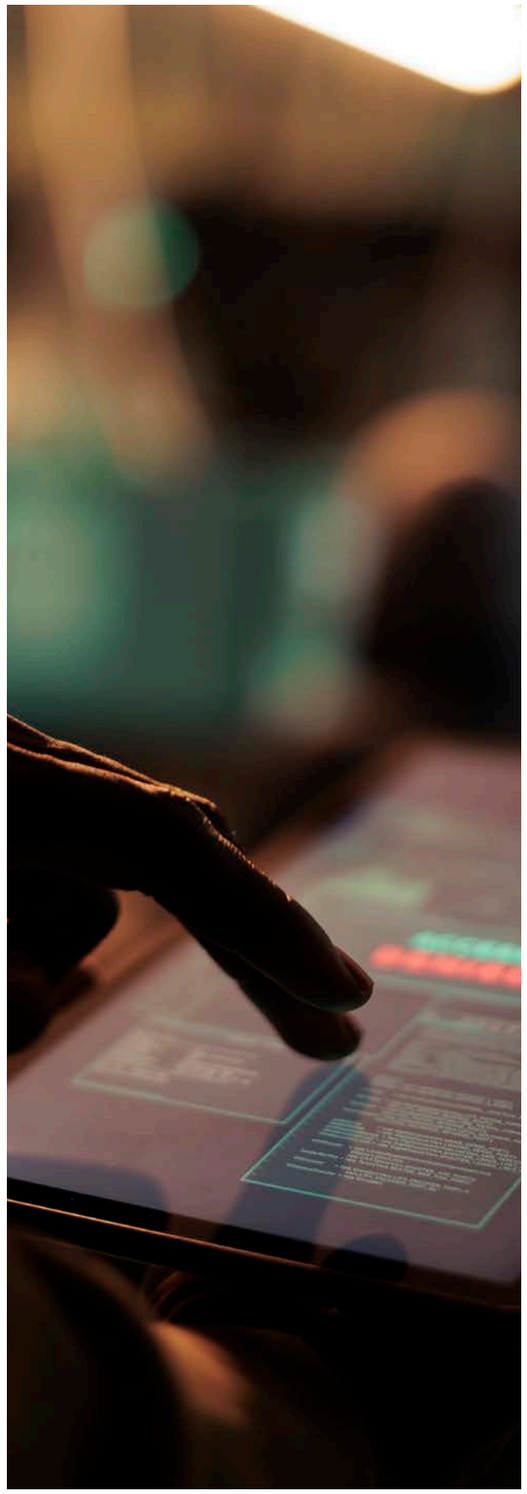
El análisis se basa en inteligencia de fuentes abiertas, informes institucionales de Europol, INTERPOL, la ONUDD, el ACNUDH y los organismos nacionales encargados de hacer cumplir la ley, investigaciones periodísticas y entrevistas a expertos. Cuando las pruebas están documentadas y se indica su fuente, se citan directamente; cuando los patrones se extraen de múltiples fuentes convergentes, se indica.

Las secciones siguientes examinan cómo la IA potencia cada uno de estos fines de reclutamiento en las principales tipologías delictivas que operan en el espacio ALC-UE.

TIPOLOGÍAS DE CAPTACIÓN CRIMINAL

Dos propósitos subyacentes impulsan el reclutamiento en todas las tipologías criminales: la adquisición de capacidades y la extracción de valor. La adquisición de capacidades consiste en captar a personas que poseen habilidades, acceso o atributos específicos que sirven a los requisitos operativos de la organización. La extracción de valor consiste en reclutar a personas que pueden ser compelidas, engañadas o inducidas a generar ingresos o a asumir riesgos legales en nombre de la organización. Estos fines no son mutuamente excluyentes: persona reclutada puede ser seleccionada por una capacidad específica y, al mismo tiempo, ser objeto de extracción de valor. En ambos propósitos, la lógica criminal subyacente es coherente. El reclutado representa una fuente de valor extraíble, y el proceso de reclutamiento está diseñado para minimizar el coste y las fricciones que conlleva su obtención.

Los grupos criminales calibran su enfoque en función de la función que están cubriendo y de la población a la que se dirigen. Las tipologías que siguen fueron seleccionadas en función de la disponibilidad de evidencia documentada y la recurrencia de patrones identificables en el espacio birregional.



CENTROS DE ESTAFAS

Los trabajadores captados en centros de estafa (*scam centres*) reciben instrucciones para llevar a cabo operaciones fraudulentas, típicamente estafas financieras, estafas románticas, esquemas de inversión en criptomonedas, extorsión y/o sextorsión.[2] Los centros de estafa son muy diversos en cuanto a su presencia física, lógica organizativa y modelos económicos.[3] Las operaciones van desde pequeñas instalaciones discretas en apartamentos, casas, hoteles y modestos espacios de oficina que priorizan la movilidad y el ocultamiento, hasta grandes complejos estructurados con departamentos dedicados a recursos humanos, TI, captación de clientes y gestión. Se ha documentado la existencia de centros de estafa que operan dentro de instalaciones penitenciarias: en Colombia, se han identificado operaciones en las prisiones de La Picota y La Modelo en Bogotá, así como en instalaciones en Tunja, Ibagué, Cali y Medellín;[4] en Bolivia, el Clan San Roque, operando dentro del Penal San Roque en Chuquisaca, Sucre, desarrolló estafas impulsadas por IA utilizando teléfonos inteligentes.

Muchos centros de fraude operan como parte de estructuras en red más amplias con alcance internacional. Un centro ubicado en un país puede mantener componentes funcionales (como infraestructura informática o procesamiento de pagos) en otro, formando lo que puede describirse como centros de fraude virtuales dispersos con nodos locales.[5]

La magnitud de la mano de obra implicada es considerable. Se ha traficado con personas procedentes de 80 países, entre ellos de América Latina y el Caribe (ALC) y la UE, para

2 Conrad, C., Heintz, E. y Mace, M. (2025). Un posible nexo entre las estafas forzadas y la sextorsión de menores con fines económicos. International Justice Mission.
3 Amerhauser, K. y Goodwin, A. (marzo de 2026). Un mundo de engaños: análisis del panorama del fenómeno de los centros de estafa globales. Iniciativa Global contra el Crimen Organizado Transnacional.
4 Amerhauser y Goodwin (2026), Un mundo de engaños.
5 Amerhauser y Goodwin (2026), Un mundo de engaños.

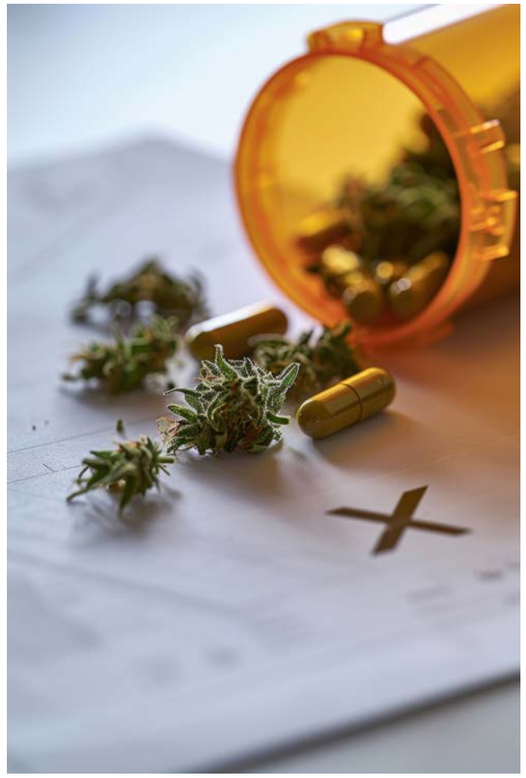
Figura 1. Ubicaciones documentadas de centros de estafa en Europa

Figura 2. Ubicaciones documentadas de centros de estafa en ALC

incorporarlas a centros de estafas en línea; se estima que solo en el sudeste asiático hay 300 000 o más personas retenidas en operaciones de estafa.[6] El reclutamiento para estas operaciones es predominantemente engañoso, y muchas víctimas denuncian haber sido reclutadas a través de aplicaciones de mensajería y plataformas de redes sociales, creyendo que iban a desempeñar un empleo legítimo en los sectores del telemarketing o la tecnología.[7] Los casos documentados abarcan países de Europa Occidental y Oriental,[8] los Balcanes Occidentales, y toda América Latina y el Caribe.

Las figuras 1 y 2 presentan casos documentados extraídos de informes de fuentes abiertas y divulgaciones de las fuerzas de seguridad que muestran los países donde se han reportado operaciones de centros de estafa.

6 Oficina de las Naciones Unidas contra la Droga y el Delito. (abril de 2025). Punto de inflexión: implicaciones globales de los centros de estafa, la banca clandestina y los mercados en línea ilícitos en el sudeste asiático.
7 INTERPOL. (2023). América: detenidos 257 presuntos traficantes de migrantes y de personas.
8 Oficina del Representante Especial y Coordinador para la Lucha contra la Trata de Seres Humanos (abril de 2026), La trata con fines de operaciones de ciberestafa: Implicaciones para la región de la OSCE — Evaluación exploratoria. Organización para la Seguridad y la Cooperación en Europa.



TRÁFICO DE DROGAS

El tráfico de drogas es el mercado delictivo con la huella de reclutamiento más amplia y diversificada tanto en la región de América Latina y el Caribe como en la UE. Más de la mitad de las redes delictivas más amenazantes que operan en la UE están involucradas en el tráfico de drogas, y estas redes reclutan cada vez más a jóvenes y personas vulnerables para absorber el riesgo legal, ampliar el alcance operativo y aislar a los actores de alto nivel de la exposición a las fuerzas del orden.[9]

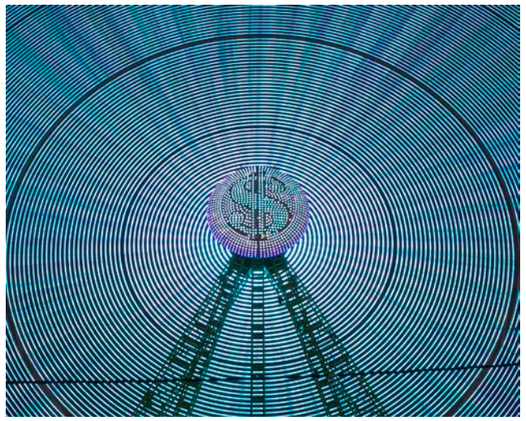
El reclutamiento para puestos como mensajeros, traficantes callejeros, vigías y recuperadores de contenedores se produce predominantemente a través de las redes sociales, donde las redes delictivas publican anuncios de empleo falsos en Instagram, Snapchat y plataformas similares, diseñados para parecerse a oportunidades de empleo legítimas.[10] En el Reino Unido, el modelo de las «county lines» ejemplifica este patrón a gran escala: las bandas urbanas reclutan a niños de tan solo siete años para transportar drogas a mercados no urbanos, con unos 14 500 niños afectados por las redes de «county lines» a partir de 2025.[11,12] En Francia, los menores reclutados a través de las redes sociales actúan como traficantes de calle desechables, mientras que en Bruselas el 65 % de las personas detenidas por tráfico de drogas eran menores.[13] En América Latina, el reclutamiento va más allá de las funciones a pie de calle para incluir funciones técnicas especializadas. Los cárteles mexicanos, en particular el *Cártel de Jalisco Nueva Generación* (CJNG) y el *Cártel de Sinaloa*, reclutan activamente a exsoldados para la operación de drones de y la fabricación de explosivos: en junio de 2025, las autoridades de seguridad mexicanas confirmaron que ambos cárteles estaban reclutando a exsoldados colombianos con experiencia en combate, tras la detención en Michoacán de 12 colombianos vinculados a un ataque con minas que causó la muerte de ocho soldados mexicanos.[14] Unidades de drones de los cárteles, como *«Los* Droneros», adscritas al *Cártel de Jalisco Nueva Generación*, llevan a cabo actualmente más de 1000 incursiones transfronterizas en Estados Unidos al mes, transportando drogas y realizando reconocimientos contra las fuerzas del orden.[15]

9 Comisión Europea. Plan de Acción de la UE contra el tráfico de drogas (COM (2025) 744 final).

10 Comisión Europea. Plan de Acción de la UE contra el tráfico de drogas (COM (2025) 744 final).
11 Parlamento del Reino Unido. (24 de febrero de 2025). Tráfico de drogas en las «county lines» [debate en el Hansard].
12 Railway Children. (17 de agosto de 2023). El rostro cambiante de las «county lines»: por qué es importante y qué estamos haciendo al respecto.
13 Iniciativa Global contra la Delincuencia Organizada Transnacional. (26 de junio de 2025). Niños soldados en Europa: ¿Por qué la delincuencia organizada recluta cada vez más a menores?
14 Associated Press. (11 de junio de 2025). El jefe de seguridad de México afirma que los cárteles de la droga están reclutando a exsoldados colombianos.
15 Felbab-Brown, V. (18 de febrero de 2026). Cómo utilizan los cárteles mexicanos los drones, ahora y en el futuro.

BLANQUEO DE CAPITALES Y MULAS FINANCIERAS

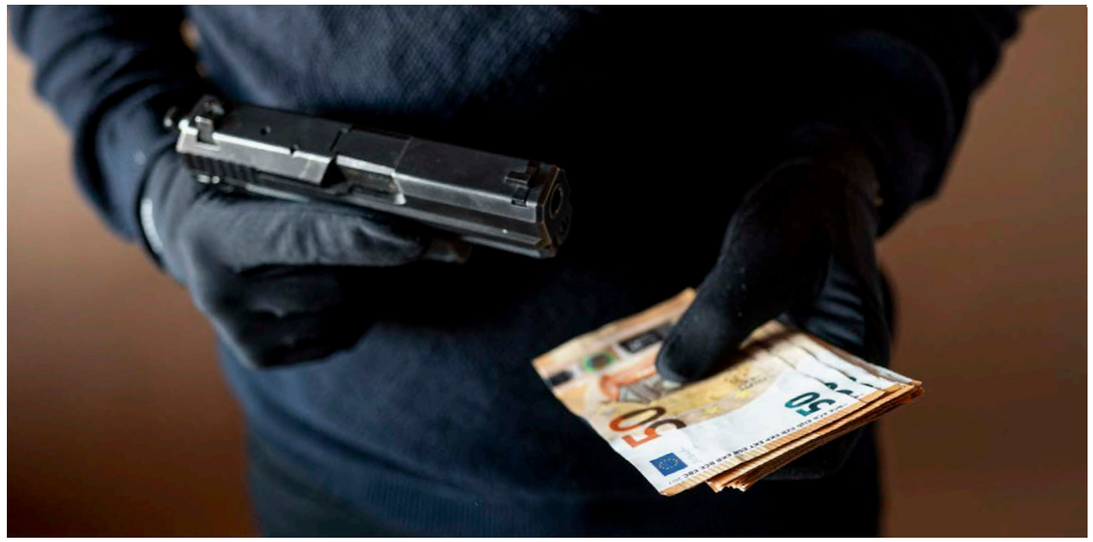
Las mulas de dinero (money mules) son personas captadas para recibir y transferir fondos ilícitos a través de sus propias cuentas bancarias o monederos de criptomonedas, proporcionando una capa de distancia entre las ganancias delictivas y su origen. Son un facilitador fundamental del crimen organizado. Sin redes de muleros que estratifiquen y muevan los fondos, la infraestructura financiera de las empresas criminales estaría significativamente más expuesta a la detección. El reclutamiento en redes de muleros financieros sigue la misma lógica que otras formas de captación criminal: los objetivos son identificados en función de su vulnerabilidad financiera, engañados mediante ofertas de empleo falsas que prometen ingresos fáciles, y a menudo ignorantes de que están participando en el blanqueo de capitales hasta después de los hechos.[16] Según la Evaluación de la Amenaza de la Delincuencia Organizada Grave de la UE (SOCTA) de 2025 de Europol, los jóvenes y las personas vulnerables son captados cada vez más como mulas financieras a través de plataformas de redes sociales y videojuegos, con redes criminales que se dirigen deliberadamente a personas con dificultades económicas, jóvenes desempleados y estudiantes.[17] Las mulas pueden ser captadas mediante engaño o incentivo financiero, incluyendo ser obligados a abrir cuentas bancarias o monederos de criptomonedas y gestionar el flujo de ganancias criminales bajo coacción. El uso deliberado de reclutas jóvenes también cumple un propósito operativo: los menores y los delincuentes primarios tienen una menor exposición legal, lo que reduce el riesgo para las redes criminales que los dirigen.[18]



La magnitud del problema es considerable. Las operaciones de la Acción Europea contra las Mulas Financieras (EMMA, por sus siglas en inglés) de Europol, realizadas anualmente en coordinación con fuerzas de seguridad de 30 o más países y más de 2.800 bancos e instituciones financieras, han identificado sistemáticamente miles de mulas por ciclo: solo la EMMA 19 identificó 10.759 mulas financieras y 474 captadores, resultando en 1.013 detenciones.[19] Las operaciones en el marco de EMMA se han extendido más allá de Europa hacia Colombia, Singapur y Australia, lo que refleja el carácter transnacional de las redes de captación de mulas.[20] Las criptomonedas han ampliado el alcance del muling más allá de las transferencias bancarias tradicionales: los denominados cibermulas mueven ahora fondos a través de monederos de activos digitales, aprovechando la velocidad y el relativo anonimato de las transacciones de criptomonedas, y el uso criminal de criptomonedas para el blanqueo de capitales se ha expandido desde la ciberdelincuencia hacia el tráfico de drogas, el tráfico de migrantes y otros mercados criminales tradicionales.[21]

16 Europol. (2025a). El ADN cambiante de la delincuencia grave y organizada: Evaluación de la amenaza de la delincuencia grave y organizada en la UE 2025 (EU-SOCTA 2025).
17 Europol (2025a), El ADN cambiante de la delincuencia grave y organizada.
18 Europol (2025a), El ADN cambiante de la delincuencia grave y organizada.

19 Europol. (2024a, julio). La operación europea contra los «mules» de dinero da lugar a 1 803 detenciones.
20 Europol. (2021). European Money Mule Action leads to 1,803 arrests.
21 Europol (2025a), El ADN cambiante de la delincuencia grave y organizada.



VIOLENCIA COMO SERVICIO Y ASESINATOS POR ENCARGO

La Violencia como Servicio (VaaS, por sus siglas en inglés) hace referencia a la externalización de actos violentos por parte de redes criminales a proveedores de servicios de terceros, frecuentemente captados en línea. En lugar de mantener una plantilla violenta permanente, los grupos criminales contratan actos que van desde la intimidación y la tortura hasta los asesinatos por encargo, pagando por cada transacción.

El Grupo Operativo GRIMM de Europol, lanzado en abril de 2025 y liderado por la Policía sueca con coordinación en once países, fue creado específicamente para abordar este fenómeno. En sus primeros doce meses de operación, el grupo de trabajo identificó a más de 1.400 personas vinculadas a redes VaaS y realizó aproximadamente 280 detenciones.[22] Actualmente, el Grupo Operativo GRIMM se ha convertido en un grupo de trabajo multinacional en el que participan más de 10 países, bajo la coordinación de Europol.

Una característica definitoria del reclutamiento para VaaS es la captación deliberada de personas jóvenes e inexpertas, incluidos menores, a través de plataformas de redes sociales, aplicaciones de mensajería cifrada y sitios de videojuegos, utilizando promesas de dinero, estatus o pertenencia a un grupo.[23]

En Dinamarca, personas de tan solo 14 años fueron detenidas en relación con intentos de asesinato ordenados a través de plataformas cifradas, con sospechosos ubicados en Suecia, Marruecos y Dinamarca.[24] Los menores son captados intencionadamente. Las redes criminales explotan deliberadamente su menor exposición legal y la percepción de impunidad procesal.[25]

22 Europol. (2026, abril). Primer aniversario de la Operación Taskforce GRIMM: objetivos de violencia-como-servicio publicados en la lista de los más buscados de la UE. Agencia de la Unión Europea para la Cooperación Policial; véase también Europol. (2025c, diciembre). Grupo de trabajo operativo GRIMM: 193 detenciones *en 6 meses en la lucha contra las redes de «violencia como servicio»*.

23 Europol. (2025d). Grupo de trabajo operativo GRIMM: Lucha contra *la violencia como servicio y el reclutamiento de jóvenes delincuentes para la delincuencia grave y organizada*.

24 Europol. (1 de julio de 2025f). Adolescentes reclutados como sicarios: Dinamarca y Suecia contraatacan la violencia como servicio.

25 Oficina Federal de Investigación. (2025, 23 de julio). The Com: El robo, la extorsión y la violencia son una amenaza creciente para los jóvenes en Internet (Alerta n.º I-072325-3-PSA). Centro de Denuncias de Delitos en Internet.

Este patrón no se limita a Europa. En América Latina, grupos incluido el Cártel de Jalisco Nueva Generación en México y el ELN en Colombia y Venezuela realizan reclutamiento forzado bajo amenazas directas, obligando a personas a asumir roles violentos mediante coacción más que mediante engaño.[26] Los cárteles mexicanos, especialmente el Cártel de Jalisco Nueva Generación, utilizan plataformas de redes sociales para reclutar a personas para roles violentos, incluidos vigías, conductores y sicarios. Una investigación del Seminario sobre Violencia y Paz de El Colegio de México encontró que, de 100 cuentas de TikTok vinculadas a la actividad de grupos criminales, el 55% estaban asociadas con el Cártel de Jalisco Nueva Generación.[27] El contenido de captación circula utilizando lenguaje codificado, emojis e insignias de cárteles para evadir la moderación de contenidos, con ofertas de trabajo publicadas en las secciones de comentarios en lugar de en los títulos de los vídeos para evitar la detección.[28]

Los cárteles también han extendido su captación digital a las plataformas de videojuegos: en México, niños de tan solo 11 años han sido contactados a través de juegos como Free Fire y Grand Theft Auto V, donde captadores que se hacen pasar por compañeros de juego invitan a los jugadores a unirse a organizaciones criminales.[29] La cadena de captación en estos casos reproduce de cerca los patrones identificados por Europol en el contexto europeo: contacto en línea, *grooming* gradual y eventual incorporación a un rol criminal, con el captador operando típicamente desde una ubicación diferente a la del perpetrador.

EXPLOTACIÓN SEXUAL

La trata con fines de explotación sexual sigue siendo el propósito más documentado por el que las mujeres y las niñas son captadas por redes criminales, tanto a nivel mundial como dentro de ALC-UE. A nivel mundial, las mujeres y las niñas constituyeron el 61% de todas las víctimas de trata detectadas en 2022, con el 60% de las niñas víctimas detectadas traficadas específicamente para la explotación sexual.[30] Dentro de la Unión Europea, los datos más recientes de Eurostat confirman que la explotación sexual representó el 46% de todos los casos de trata registrados en 2024, con mujeres y niñas representando el 63% de las 9.678 víctimas registradas.[31] Ambos conjuntos de datos reflejan únicamente a las víctimas detectadas y registradas, y se comprende ampliamente que subestiman la magnitud real del delito. La captación para la explotación sexual sigue un patrón consistente de engaño: las víctimas son abordadas mediante ofertas de empleo fabricadas, oportunidades falsas de modelaje o trabajo doméstico, o relaciones románticas manufacturadas, y posteriormente controladas mediante servidumbre por deudas, confiscación de documentos, amenazas y violencia.

En el corredor ALC-UE específicamente, INTERPOL identificó en 2024 un flujo de trata estructurado y creciente en el que las víctimas son atraídas desde América Latina mediante ofertas de empleo engañosas y transportadas a Europa para su explotación mediante prostitución forzada.[32] El informe de la UNODC de 2024 señala además que las redes de crimen organizado tienen un impacto notablemente mayor sobre el número de víctimas y el alcance geográfico que los traficantes individuales no organizados, y que las mujeres actúan como captadoras

26 EL PACCTO 2.0. (2025, septiembre). El uso de la inteligencia artificial por parte de redes criminales de alto riesgo. Unión Europea / Expertise France. DOI 10.5281/zenodo.16750778.

27 Seminario sobre Violencia y Paz, El Colegio de México. (2025). Nuevas fronteras para el reclutamiento digital [citado en TechTimes, 4 de diciembre de 2025].

28 Bledsoe, R. (28 de mayo de 2025). El papel de las redes sociales en el reclutamiento de los cárteles.

29 Dalby, C. (15 de octubre de 2021). Cómo utilizan los cárteles mexicanos los videojuegos para reclutar a niños.

30 Oficina de las Naciones Unidas contra la Droga y el Delito. (diciembre de 2024). Informe mundial sobre la trata de personas 2024.

31 Eurostat. (29 de enero de 2026). 9 678 víctimas de trata de personas registradas en 2024. Comisión Europea.

32 INTERPOL. (18 de diciembre de 2024). Detalles de la investigación de INTERPOL sobre la trata de personas facilitada por Internet.

en una proporción significativa de los casos, aprovechando las relaciones de confianza para entregar a las víctimas a la explotación.[33] Una tendencia adicional y notable dentro de la UE es el aumento sostenido de la trata con fines de trabajo forzado, que alcanzó el 37% de los casos registrados en 2024, reduciendo la brecha histórica con la explotación sexual y reflejando la creciente diversificación de los propósitos de captación criminal.[34]

EXPLOTACIÓN Y ABUSO SEXUAL DE MENORES EN LÍNEA

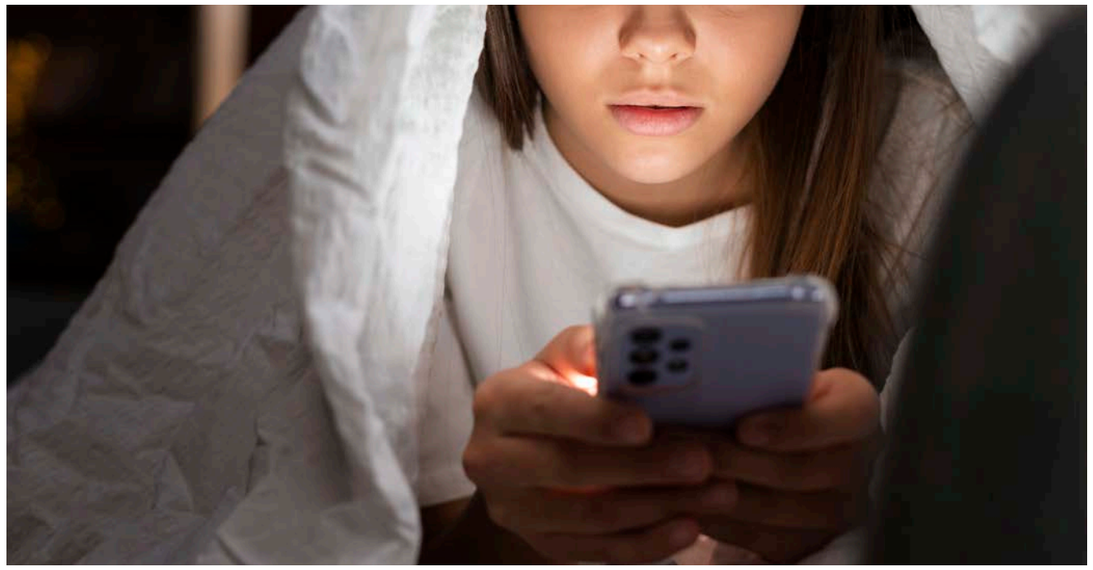
La explotación y el abuso sexual infantil en línea (OCSEA, por sus siglas en inglés) y la producción de material de abuso sexual infantil (CSAM) han ido adquiriendo cada vez más las características del crimen organizado. Lo que antes estaba impulsado principalmente por delincuentes individuales motivados por la gratificación sexual ha cambiado: ahora son grupos delictivos organizados los que explotan a los niños con fines lucrativos, y algunos de sus miembros no tienen ningún interés personal en el abuso en sí.[35] Algunas operaciones de CSAM a gran escala reflejan la estructura de las empresas criminales, con roles definidos que incluyen reclutadores, facilitadores y gestores de pagos, y operan a través de las fronteras utilizando comunicaciones cifradas, plataformas de la web oscura y criptomonedas para evadir la detección.[36] La plataforma Kidflix, cerrada en marzo de 2025 como parte de la Operación Stream liderada por Europol en la que participaron 38 países, ilustraba esta sofisticación organizativa: contaba con 1,8 millones de usuarios, 91 000 vídeos de abuso y gestionaba una economía de tokens que recompensaba a los usuarios por subir material nuevo.[37] El CSAM también se utiliza como mecanismo de selección en redes delictivas más amplias: a los posibles miembros de grupos cerrados en plataformas como Telegram se les exige a veces descargar CSAM en sus dispositivos como prueba de compromiso.[38]

El reclutamiento en estas redes explota tanto a niños como a adolescentes mayores mediante una serie de métodos engañosos y coercitivos. Las redes organizadas llevan a cabo el reclutamiento fuera de línea a través de ofertas falsas de trabajo o de modelaje antes de coaccionar a las víctimas para que participen en abusos en línea.[39] En línea, se captan a los niños a través de las redes sociales y de salas de chat ocultas en aplicaciones como Telegram y WeChat. Los adolescentes mayores (en particular los chicos de entre 14 y 17 años) son blanco de extorsión con coacción sexual, que ha aumentado considerablemente. Desde 2022, las denuncias al servicio «Report Remove» de la IWF relacionadas con chicos de entre 14 y 17 años se han multiplicado por trece, mientras que el número de imágenes y vídeos relacionados se ha multiplicado por once.[40] La IWF registró 63 044 niños de entre 14 y 17 años en imágenes de abuso sexual delictivo solo en 2025, repartidos en 56 179 imágenes.[41]

Abuso de niños en redes extremistas violentas nihilistas

Una categoría diferenciada y creciente de captación y abuso de menores opera a través de redes extremistas violentas nihilistas (NVE), de las cuales 764 y su precursora

33 ONUDC (2024), Informe mundial sobre la trata de personas 2024.
34 Eurostat. (2026). Estadísticas sobre la trata de seres humanos. Comisión Europea.
35 Europol (2025a), El ADN cambiante de la delincuencia grave y organizada: Evaluación de la amenaza de la delincuencia grave y organizada en la UE 2025 (EU-SOCTA 2025). Agencia de la Unión Europea para la Cooperación Policial.
36 Europol. (2025a). El ADN cambiante de la delincuencia grave y organizada.; Europol. (2024b). Evaluación de la amenaza de la delincuencia organizada en Internet (IOCTA) 2024. Agencia de la Unión Europea para la Cooperación Policial.
37 Europol. (2 de abril de 2025b). Operación mundial contra Kidflix, una importante plataforma de explotación sexual infantil con casi dos millones de usuarios [Comunicado de prensa].
38 Comolli, V. (2025). Explotación sexual infantil organizada: abordar los motivos y las necesidades de respuesta en el sudeste asiático. Iniciativa Global contra la Delincuencia Organizada Transnacional.
39 Comolli (2025), Explotación sexual infantil organizada.
40 Fundación Internet Watch. (2025). Tomar medidas para ayudar a los adolescentes mayores. Informe anual de datos y análisis de la IWF 2025.
41 IWF (2025), Tomar medidas para ayudar a los adolescentes mayores.



«CVLT»[42] son las más documentadas. Estas redes captan a niños y adolescentes, principalmente de entre 10 y 17 años, a través de plataformas de redes sociales, aplicaciones de videojuegos y servicios de mensajería cifrada, dirigiéndose deliberadamente a personas con vulnerabilidades preexistentes como depresión, trastornos alimentarios y aislamiento social.[43] Los reclutadores establecen relaciones de confianza o relaciones románticas simuladas antes de pasar a la coacción, chantajeando a las víctimas para que produzcan material de abuso sexual infantil, cometan actos de autolesión, hagan daño a animales y, en algunos casos, intenten suicidarse o ejerzan violencia contra otros, con el contenido retransmitido en directo y difundido dentro de la red como una forma de moneda de cambio y estatus.[44] La red es descentralizada pero está organizada: 764 surgió en 2021 a partir de «CVLT» y opera a través de una estructura jerárquica de administradores, reclutadores y autores en múltiples jurisdicciones.[45] En diciembre de 2025, el FBI llevaba a cabo activamente más de 350 investigaciones relacionadas con «764» y grupos NVE afines en sus 55 oficinas locales, y se estimaba que el número de víctimas conocidas ascendía a miles en todo el mundo.[46] Canadá ha designado a 764 como entidad terrorista en virtud de su Código Penal; el Departamento de Justicia de Estados Unidos la ha descrito como una de las organizaciones de explotación infantil más graves con las que se ha encontrado jamás.[47] Las acusaciones federales relacionadas con 764 aumentaron un 350 % entre 2024 y 2025, y el 59 % de los cargos se referían a delitos relacionados con material de abuso sexual infantil (CSAM).[48]

42 CVLT: red de sextorsión denominada «CVLT»," que obligaba a menores a realizar actos sexuales ante la cámara y utilizaba el material obtenido para seguir extorsionándolos. Fuente: ADL (2025) "Backgrounder 764", 20 de junio de 2025.
43 Oficina Federal de Investigaciones. (marzo de 2025). Las redes violentas en línea se dirigen a poblaciones vulnerables y menores de edad en todo Estados Unidos y en todo el mundo (Alerta PSA I-030625-PSA). Centro de Denuncias de Delitos en Internet.
44 Departamento de Justicia de EE. UU. (30 de octubre de 2025a). Líder de Arizona de la red extremista violenta «764» acusado de dirigir una empresa de explotación infantil. Oficina de Asuntos Públicos.
45 Centro Nacional de Innovación, Tecnología y Educación Antiterrorista. (Agosto de 2025). El juicio contra «764»: un análisis de los cargos federales y estatales. Universidad de Nebraska-Omaha.
46 Departamento de Justicia de EE. UU. (2025b, diciembre). El líder del grupo extremista «764» se declara culpable de los cargos de la ley RICO y de explotación infantil. Oficina de Asuntos Públicos.
47 NCITE (2025), El juicio contra 764.
48 NCITE (2025), El procesamiento de 764.

EL USO DE LA IA EN EL RECLUTAMIENTO CRIMINAL

La IA está amplificando los modelos tradicionales de captación, aumentando la velocidad, la escala, la credibilidad y la resiliencia operativa en todo el ecosistema de explotación criminal.[49] La IA no cambia la lógica subyacente del reclutamiento criminal. Los grupos criminales aún necesitan llegar a posibles reclutas, identificar a los objetivos más viables, captarlos de manera convincente y evitar la detección mientras lo hacen. Lo que cambia la IA es el coste, la escala y la precisión con que pueden realizarse estas funciones. Las barreras que anteriormente limitaban el reclutamiento criminal, como el idioma, la geografía, la capacidad de producción y la necesidad de operadores especializados, se han reducido o eliminado de forma sustancial. El resultado es una versión acelerada e industrializada de los patrones de captación existentes. Las cinco funciones siguientes proporcionan un modelo para describir cómo la IA amplifica cada etapa del proceso de reclutamiento.

49 Montes Delijorge, J. (30 de abril de 2026). La IA está transformando el panorama de la explotación humana: un análisis de las amenazas en evolución. Medium.

Figura 3. Las cinco funciones de la IA en el reclutamiento criminal

ALCANCE

Las herramientas de IA permiten a las organizaciones criminales producir y distribuir contenido de captación a una escala y calidad que antes estaban fuera de su alcance, eliminando simultáneamente las limitaciones del idioma, la capacidad de producción y la presencia local. Las plataformas de redes sociales funcionan como infraestructura lista para usar: sus algoritmos de recomendación, sistemas publicitarios y herramientas de segmentación de audiencias son explotados por los actores criminales en lugar de ser construidos por ellos.

- **Acceso ampliado a idiomas.** Las herramientas de traducción de IA producen texto fluido en cientos de idiomas, incluidos dialectos informales y

jergas de subculturas, lo que permite a las organizaciones criminales comunicarse de manera convincente dentro de comunidades que antes estaban fuera de su alcance.[50,51,52,53]

- **Captación demográfica más amplia.** Las plataformas de redes sociales ofrecen segmentación integrada, lo que permite a los actores criminales explotar algoritmos de recomendación que muestran contenido a los usuarios basándose en datos de perfil y comportamiento de búsqueda, llegando a personas con dificultades financieras, jóvenes desempleados y otros grupos vulnerables sin necesidad de infraestructura técnica personalizada.[54,55]

- **Mayor diversidad de formatos de contenido.** La IA permite a las organizaciones criminales producir contenido atractivo en múltiples formatos (incluidos vídeos cortos, música y material narrativo) que normaliza los mensajes de captación y amplía el alcance de la audiencia.[56]

- **Contenido de calidad profesional a bajo coste.** La IA permite a las organizaciones criminales producir materiales de captación pulidos, incluidas ofertas de trabajo, páginas web y documentación de apoyo, que cumplen los estándares de apariencia de los empleadores legítimos.[57]

50 Organización para la Seguridad y la Cooperación en Europa y Oficina de Apoyo Regional del Proceso de Bali. (2024). Nuevas fronteras: el uso de la inteligencia artificial generativa para facilitar la trata de personas. OSCE.
51 Aldawsari, A. H. (2024). Evaluación de herramientas de traducción: Google Translate, Bing Translator y Bing AI en coloquialismos árabes. Arab World English Journal (AWEJ), número especial sobre ChatGPT, 237–251.
52 Amerhauser y Goodwin (2026), Un mundo de engaños.
53 McPherson, P. (15 de septiembre de 2025). ChatGPT se utilizó «para ayudar a los estafadores a hacer de las suyas» en un complejo de fraude en Asia. Reuters.

54 Velasco, C., García Periche, J., Gómez Gómez, J. D., y Bueno Benedí, M. (2025). Inteligencia artificial y delincuencia organizada (informe EL PACCTO 2.0). Expertise France y la Fundación para la Internacionalización de las Administraciones Públicas. https://doi.org/10.5281/zenodo.16740421
55 Europol (2025a), El ADN cambiante de la delincuencia grave y organizada.
56 Velasco et al. (2025), Inteligencia artificial y delincuencia organizada.
57 OSCE y RSO del Proceso de Bali (2024), Nuevas fronteras; Levesque (2025), La amenaza del tráfico potenciada por la IA.

Figura 4. Ejemplo de publicidad generada por IA con indicios de posible explotación laboral

El anuncio, distribuido en WhatsApp, está generado por IA. Presenta un aspecto profesional, utiliza un inglés fluido y no contiene errores lingüísticos ni gramaticales. Los iconos, el formato y la maquetación tienen un aspecto profesional. Hay indicios de posible explotación laboral. Incluye ofertas de comida, alojamiento y transporte gratuitos, así como un «cargo por servicio» que debe pagar el candidato o su agente.

IDENTIFICACIÓN

Una vez que el contenido de captación llega a una amplia audiencia, las herramientas de IA permiten a las organizaciones criminales pasar de la difusión masiva a la captación precisa, identificando a las personas más susceptibles de ser engañadas o coaccionadas. Este cambio de un alcance amplio a una selección quirúrgica aumenta significativamente la eficiencia y efectividad del reclutamiento criminal.[58,59]

- **Análisis de datos multimodal.** Las herramientas impulsadas por IA pueden recopilar y analizar sistemáticamente datos públicamente disponibles de redes sociales, aplicaciones de citas y plataformas de mensajería, procesando simultáneamente texto, imágenes, vídeo, datos de ubicación y conexiones de redes sociales para construir perfiles detallados de los posibles objetivos.[60,61] Los grandes conjuntos de datos con información personal también pueden recopilarse fácilmente, así como comprarse en línea y en la dark web.[62]

- **Detección de vulnerabilidades.** Mediante el Procesamiento del Lenguaje Natural, la IA puede detectar indicadores de dificultades financieras, aislamiento social y vulnerabilidad emocional a través de la elección de palabras, el sentimiento y los patrones de comunicación, identificando a las personas de alto riesgo.[63,64] El aumento en la detección de vulnerabilidades puede dar lugar a nuevas metodologías de captación y a nuevos tipos de víctimas.

- **Refinamiento progresivo de la captación**. Los sistemas de captación habilitados por IA pueden configurarse para refinar los criterios de captación a lo largo del tiempo cuando los operadores recopilan datos de resultados y los retroalimentan al sistema.[65] En tales

58 Levesque, J. (2025). La amenaza del tráfico potenciada por la IA: Examen de la evolución tecnológica de las operaciones de trata de personas con herramientas de IA.
59 OSCE y RSO del Proceso de Bali (2024), Nuevas fronteras.

60 Levesque, J. (2025). La amenaza del tráfico potenciada por la IA: Examen de la evolución tecnológica de las operaciones de trata de personas con herramientas de IA.
61 Srinath, S., y Srivastava, S. (2024). Rastreo web dinámico potenciado por IA generativa en Azure mediante Automation Account y GPT-3.5. Semantic Scholar.
62 Amerhauser y Goodwin (2026), Un mundo de engaños.
63 Levesque, J. (2025). La amenaza del tráfico potenciada por la IA: Examen de la evolución tecnológica de las operaciones de trata de personas con herramientas de IA.
64 Bender, E. M., Gebru, T., McMillan-Major, A., y Shmitchell, S. (2021). Sobre los peligros de los loros estocásticos: ¿pueden los modelos de lenguaje ser demasiado grandes? En Actas de la Conferencia ACM 2021 sobre Equidad, Responsabilidad y Transparencia (pp. 610–623). ACM. https://doi.org/10.1145/3442188.3445922
65 Chaudhari, S., Aggarwal, P., Murahari, V., Rajpurohit, T., Kalyan, A., Narasimhan, K., & Castro da Silva, B. (2025). RLHF deciphered: A critical analysis of reinforcement learning from human feedback for LLMs. ACM Computing Surveys, 58(2). https://doi.org/10.1145/3743127

casos, las interacciones previas pueden utilizarse para ajustar la selección de audiencia, los mensajes, el momento y las estrategias de seguimiento. Cuando se implementan deliberadamente tales bucles de retroalimentación, los sistemas de captación criminal pueden volverse más eficientes y más difíciles de detectar con el tiempo.

INTERACCIÓN

Una vez identificado un objetivo, las herramientas de IA permiten a las organizaciones criminales iniciar y mantener el contacto de formas convincentes, personalizadas y escalables. El objetivo en esta etapa es establecer la confianza o dependencia suficientes para que el objetivo acceda a cumplir sus órdenes. La IA reduce significativamente el coste y las habilidades necesarias para hacerlo de manera eficaz, al tiempo que aumenta la sofisticación del enfoque.[66]

- **Legitimidad ficticia.** Los modelos de lenguaje grande (LLM) pueden generar ofertas de empleo profesionales, sitios web de empresas, contratos de trabajo, perfiles de empleados y documentación de apoyo que se ajustan a las normas sectoriales y nacionales, creando una apariencia convincente de legitimidad institucional.[67,68]

A fake New York State Bar Association membership card generated using our models. Redactions by OpenAI investigators.

Figura 5. Una tarjeta de miembro falsa del Colegio de Abogados del Estado de Nueva York generada utilizando modelos de OpenAI. (Censurado por los investigadores de OpenAI)

La licencia falsa se utilizó en la «Operación False Witness», que implicaba un grupo de cuentas de ChatGPT utilizadas para promocionar al menos seis bufetes de abogados falsos que ofrecían servicios de recuperación de estafas.[69] Se utiliza como ejemplo del uso de la IA para generar legitimidad institucional.

- **Comunicación personalizada.** Los modelos de lenguaje grande (LLM) generan contenido gramaticalmente correcto y adecuado al contexto, capaz de imitar el lenguaje, el tono y las referencias culturales de un objetivo, produciendo mensajes que parecen provenir de un compañero de confianza o de alguien con información privilegiada.[70]

- **Manipulación romántica y de relaciones.** La IA permite a los delincuentes iniciar y mantener relaciones románticas ficticias. Esto puede hacerse a través de operadores humanos asistidos por IA o de chatbots totalmente automatizados capaces de mantener conversaciones coherentes y emocionalmente receptivas durante semanas o meses, creando dependencia emocional en las víctimas.[71,72] La interacción con chatbots de IA se ha relacionado con un aumento de la dependencia emocional, el aislamiento social y acciones en el mundo real por parte de los usuarios, incluyendo casos que han resultado en daños graves.[73,74]

- **Coacción mediante material generado por IA.** Las herramientas de deepfake pueden generar imágenes sexuales sintéticas de una persona concreta a partir de un pequeño número de fotografías ordinarias, sin que haya existido ningún material real. Estas imágenes pueden utilizarse para coaccionar a las víctimas amenazándolas con su divulgación, lo que las obliga a obedecer o a mantener la interacción. Los deepfakes generados por IA eliminan la necesidad de obtener primero material comprometedor real, acortando el tiempo entre el contacto inicial y el control coercitivo y reduciendo la exposición legal del agresor, al tiempo que amplifican el daño a la víctima.[75,76,77]

- **Interacción autónoma y aumentada a gran escala.** Los sistemas de IA pueden gestionar cientos de conversaciones de reclutamiento simultáneas sin supervisión humana, adaptando las respuestas en función de las respuestas de cada objetivo.[78] Cuando no se implementa la automatización total, los modelos de lenguaje grande (LLM) complementan a los reclutadores humanos generando guiones adaptativos que guían las interacciones en directo, ampliando el alcance de los operadores individuales mucho más allá de lo que sería posible de otro modo.[79]

- **Preselección de víctimas asistida por IA.** Cada vez se utilizan más los chatbots de voz para establecer el contacto inicial con víctimas potenciales a escala industrial, actuando como un filtro que califica a los objetivos antes de pasarlos a los operadores humanos.[80] Este modelo reduce el coste de la mano de obra humana en las operaciones de fraude y reclutamiento, al tiempo que aumenta significativamente el número de personas con las que se puede contactar y evaluar simultáneamente.

EVADIR

Las herramientas de IA permiten a los delincuentes llevar a cabo el reclutamiento con un riesgo operativo reducido y una mayor resiliencia ante las interrupciones. Esto funciona a través de varios mecanismos.

- **Evasión de la moderación de las plataformas.** La IA permite la adaptación en tiempo real del contenido de reclutamiento para evitar la detección automatizada. Los actores delictivos pueden rotar palabras clave, modificar imágenes y sustituir el lenguaje codificado, lo que permite que las publicaciones individuales parezcan inofensivas para los sistemas de moderación de contenidos, sin dejar de ser legibles para el público

66 Levesque, J. (2025). La amenaza del tráfico potenciada por la IA: Examen de la evolución tecnológica de las operaciones de trata de personas con herramientas de IA.
67 Levesque, J. (2025). La amenaza del tráfico potenciada por la IA: Examen de la evolución tecnológica de las operaciones de trata de personas con herramientas de IA.
68 Europol, UNICRI y Trend Micro. (2020). Usos maliciosos y abusos de la inteligencia artificial. Europol.

69 OpenAI. (2026, febrero). Interrumpiendo los usos maliciosos de nuestros modelos: Actualización. OpenAI.
70 OSCE y RSO del Proceso de Bali (2024), Nuevas fronteras.

71 Levesque, J. (2025). La amenaza del tráfico potenciada por la IA: Examen de la evolución tecnológica de las operaciones de trata de personas con herramientas de IA.
72 Chu, M. D., Gerard, P., Pawar, K., Bickham, C. y Lerman, K. (2025). Ilusiones de intimidad: apego emocional y riesgos psicológicos emergentes en las relaciones entre humanos y IA. arXiv. https://doi.org/10.48550/arXiv.2505.11649
73 Horwitz, J. (14 de agosto de 2025). El chatbot de IA coqueto de Meta invitó a un jubilado a Nueva York. Nunca llegó a casa. Reuters.
74 Duffy, C. (30 de octubre de 2024). «No hay barreras de seguridad». Esta madre cree que un chatbot de IA es responsable del suicidio de su hijo. CNN.
75 Oficina Federal de Investigación. (5 de junio de 2023). Actores maliciosos que manipulan fotos y vídeos para crear contenido explícito y esquemas de sextorsión [Anuncio de servicio público]. Centro de Denuncias de Delitos en Internet.
76 FinCEN (2025). FinCEN emite un aviso sobre la sextorsión con motivaciones económicas. Departamento del Tesoro de EE. UU., Red de Control de Delitos Financieros.
77 NCMEC (2025). El NCMEC publica nuevos datos: 2024 en cifras. Centro Nacional para Niños Desaparecidos y Explotados.

78 Levesque, J. (2025). La amenaza del tráfico potenciada por la IA: Examen de la evolución tecnológica de las operaciones de trata de personas con herramientas de IA.
79 Levesque, J. (2025). La amenaza del tráfico potenciada por la IA: Examen de la evolución tecnológica de las operaciones de trata de personas con herramientas de IA.
80 Europol. (2026). Evaluación de la amenaza de la delincuencia organizada en Internet (IOCTA) 2026: Panorama de amenazas en evolución. Europol.

destinatario.[81,82] Las ofertas de empleo se colocan en las secciones de comentarios en lugar de en los pies de foto para reducir la visibilidad ante los filtros de la plataforma.[83] Cuando se elimina una cuenta, las identidades sintéticas generadas por IA permiten una rápida reconstitución de la presencia en la misma plataforma o en otras alternativas.[84]

- **Evasión de la detección a través de la calidad del contenido.** La IA elimina los marcadores lingüísticos (como errores gramaticales, frases poco naturales o formato inconsistente) que antes permitían a los destinatarios y a las plataformas identificar las comunicaciones fraudulentas. Una oferta de empleo fraudulenta generada por un modelo de lenguaje grande (LLM) puede resultar difícil de distinguir del contenido de reclutamiento legítimo basándose únicamente en el contenido.[85,86]

- **Anonimización operativa.** Las plataformas diseñadas específicamente permiten a los operadores gestionar cientos de cuentas simultáneamente, generar números de verificación aleatorios y procesar pagos, todo ello mientras se oculta la identidad del operador y se evade la detección a nivel de plataforma.[87] La compartimentación del proceso de reclutamiento entre herramientas de IA, plataformas y gestores humanos limita aún más la exposición de cualquier actor individual.

- **Ofuscación jurisdiccional y relacional.** La traducción asistida por IA y la coordinación a distancia permiten a los instigadores, reclutadores, facilitadores locales y autores permanecer desconectados social y geográficamente. La ausencia de una relación previa entre el instigador y el autor complica la atribución y reduce la exposición de los actores de alto nivel.

- **Pruebas de demanda antes de la puesta en marcha.** Se utiliza contenido sintético para probar los mensajes de captación y evaluar la demanda antes de introducir a víctimas reales en el proceso, lo que permite a las organizaciones delictivas perfeccionar su estrategia y reducir el riesgo operativo.[88] Esto reduce el coste de los fallos en la fase de captación y permite a los actores optimizar sus métodos antes de comprometer recursos.

ESCALA

La IA no solo mejora las funciones de reclutamiento individuales, sino que también permite a las organizaciones criminales llevar a cabo el reclutamiento como una operación continua a gran escala. Una vez implementada, la automatización mediante IA transforma esencialmente el reclutamiento de un proceso intensivo en mano de obra a uno industrial.[89,90]

- **Reclutamiento simultáneo.** Un único sistema de IA puede llevar a cabo cientos de conversaciones de reclutamiento en paralelo de forma simultánea, manteniendo el contexto, adaptando el tono y haciendo un seguimiento de las interacciones previas en todas ellas con una supervisión humana mínima o nula.[91,92]

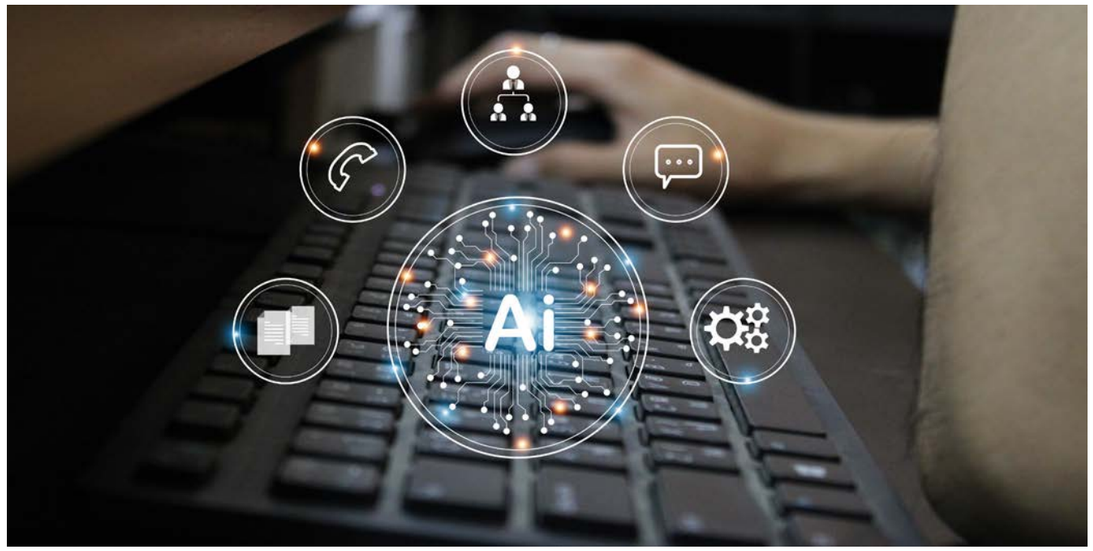
- **Sistemas de reclutamiento que se perfeccionan por sí mismos.** Una vez configurados, los sistemas de reclutamiento basados en IA sirven para refinar las tácticas de selección de objetivos, comunicación y captación basándose en los resultados de las conversaciones simultáneas, y pueden mejorar su eficacia con el tiempo sin necesidad de instrucciones humanas.[93,94] Esto significa que las operaciones de reclutamiento delictivo pueden desarrollar nuevas tácticas inesperadas a través de su propio aprendizaje operativo, lo que las hace cada vez más difíciles de identificar y contrarrestar.[95]

- **Ampliación de la estrategia.** Las herramientas de IA pueden generar, probar y adaptar estrategias de reclutamiento en diferentes contextos culturales, lingüísticos y demográficos con una profundidad que supera la capacidad humana individual.[96,97] Esto cambia no solo la eficiencia del reclutamiento criminal, sino potencialmente su carácter, ya que el desarrollo de estrategias impulsado por la IA puede producir enfoques que difieren de los patrones establecidos y que, por lo tanto, son más difíciles de reconocer para las fuerzas del orden.

- **La IA agentiva como una trayectoria de amenaza emergente.** Los actores delictivos han comenzado a aprovechar sistemas de IA agentiva capaces de llevar a cabo de forma autónoma flujos de trabajo delictivos completos.[98] Si bien el despliegue total de la IA agentiva en las operaciones de reclutamiento sigue siendo una preocupación emergente, su trayectoria apunta a que el reclutamiento delictivo se está desvinculando cada vez más de la supervisión humana, lo que dificulta significativamente la atribución y la interrupción de las actividades.

En conjunto, estas capacidades representan un cambio estructural en la economía y el ecosistema del reclutamiento criminal. La mano de obra humana era la restricción operativa determinante. La IA disuelve esa restricción. Un solo operador con acceso a herramientas comerciales puede ahora llevar a cabo lo que antes habría requerido una gran plantilla humana, multilingüe y logísticamente compleja. Esto reduce la barrera de entrada para nuevos actores criminales y permite a las redes existentes ampliar su capacidad de reclutamiento sin aumentos proporcionales en el coste, la dotación de personal o la exposición.

81 Tynan, R. (24 de abril de 2026). Cómo utilizan los cárteles mexicanos el contenido generado por IA para modernizar el reclutamiento. Alice (Medium).
82 Bledsoe, R. (28 de mayo de 2025). El papel de las redes sociales en el reclutamiento de los cárteles. Centro de Estudios Estratégicos e Internacionales.
83 Bledsoe (2025), El papel de las redes sociales en el reclutamiento de los cárteles.
84 Velasco et al. (2025), Inteligencia artificial y delincuencia organizada.
85 Velasco et al. (2025), Inteligencia artificial y delincuencia organizada.
86 OSCE y RSO del Proceso de Bali (2024), Nuevas fronteras.
87 Di Girolamo, M. (2026). Engañoso por diseño: Las herramientas basadas en IA que alimentan la industria de las estafas. C4ADS.
88 Montes Delijorge (2026), La IA está transformando el panorama de la explotación humana.
89 Levesque, J. (2025). La amenaza del tráfico potenciada por la IA: Examen de la evolución tecnológica de las operaciones de trata de personas con herramientas de IA.
90 Montes Delijorge (2026), La IA está transformando el panorama de la explotación humana.
91 Levesque, J. (2025). La amenaza del tráfico potenciada por la IA: Examen de la evolución tecnológica de las operaciones de trata de personas con herramientas de IA.
92 OpenAI. (2024). Informe técnico de GPT-4. OpenAI.
93 Levesque, J. (2025). La amenaza del tráfico potenciada por la IA: Examen de la evolución tecnológica de las operaciones de trata de personas con herramientas de IA.
94 Chaudhari et al. (2025), RLHF descifrado.
95 Velasco et al. (2025), Inteligencia artificial y delincuencia organizada.
96 Levesque, J. (2025). La amenaza del tráfico potenciada por la IA: Examen de la evolución tecnológica de las operaciones de trata de personas con herramientas de IA.
97 Departamento de Seguridad Nacional de EE. UU. (2024). Impacto de la inteligencia artificial (IA) en las actividades delictivas e ilícitas. Programa de Intercambio Analítico Público-Privado.
98 Europol (2026), IOCTA 2026.



HERRAMIENTAS Y TECNOLOGÍAS DE IA UTILIZADAS EN EL RECLUTAMIENTO

Las herramientas de IA utilizadas en el reclutamiento criminal se dividen en tres categorías: herramientas comerciales ampliamente disponibles reutilizadas con fines criminales; aplicaciones que producen contenido prohibido por la ley o por las políticas del proveedor pero derivadas de modelos comerciales mediante técnicas de *jailbreak*; y herramientas diseñadas y operadas específicamente para facilitar actividades criminales.

HERRAMIENTAS DE IA DE DOBLE USO

Para la mayoría de los fines de reclutamiento delictivo, no se requieren herramientas especializadas. Los grandes modelos de lenguaje (LLM), incluidos los modelos comerciales de vanguardia de OpenAI, Anthropic, Google, así como ByteDance y otros, proporcionan la mayor parte de las capacidades que necesitan las operaciones de reclutamiento delictivo, como redactar anuncios de empleo, generar mensajes de contacto personalizados, traducir comunicaciones, producir medios sintéticos y mantener conversaciones semiautomatizadas a gran escala.[99,100,101] La intención maliciosa es contextual, no está contenida en el contenido. Por ejemplo, una oferta de empleo fraudulenta es técnicamente indistinguible de una legítima y, por lo general, no activa los sistemas de moderación de contenidos, por lo que no requiere ninguna modificación del modelo subyacente.[102,103] Las herramientas relevantes no son necesariamente plataformas de IA delictivas creadas específicamente para ese fin, sino herramientas comunes de traducción, generación de texto, generación de imágenes, generación de música y automatización de la interacción que se utilizan para salvar las barreras lingüísticas,

99 Velasco et al. (2025), Inteligencia artificial y delincuencia organizada.
100 Europol (2025a), El ADN cambiante de la delincuencia grave y organizada.
101 Levesque, J. (2025). La amenaza del tráfico potenciada por la IA: Examen de la evolución tecnológica de las operaciones de trata de personas con herramientas de IA.
102 Velasco et al. (2025), Inteligencia artificial y delincuencia organizada.
103 OSCE y RSO del Proceso de Bali (2024), Nuevas fronteras.

crear contenido sobre estilos de vida delictivos, inflar la prueba social y mantener el contacto a través de las fronteras.

MODELOS DE CÓDIGO ABIERTO AUTOALOJADOS Y MODIFICADOS

Una categoría distinta de riesgo surge de los LLM de código abierto que se descargan, se ejecutan localmente y se operan por completo fuera de la supervisión del proveedor. A diferencia de los modelos comerciales a los que se accede a través de una API, los modelos desplegados localmente se rigen únicamente por las intenciones de sus operadores: ningún proveedor puede suspender el acceso, imponer filtros de contenido ni conservar registros de uso.

Los modelos de código abierto, como la serie LLaMA de Meta, Mistral, Falcon y las variantes de GPT-J, están disponibles públicamente para su descarga e implementación local, normalmente a través de marcos como Ollama, que permiten a los usuarios ejecutar LLM en hardware personal con unos requisitos técnicos mínimos.[104] Las variantes sin censura (modelos desprovistos de mecanismos de seguridad) se anuncian abiertamente en foros delictivos y son fáciles de localizar. Cisco Talos documentó modelos sin censura, como OnionGPT y WhiteRabbitNeo, anunciados en foros de piratería informática, con Llama 2 Uncensored disponible para su implementación local a través de Ollama y capaz de generar contenido de phishing y otros resultados dañinos sin restricciones.[105]

La magnitud de la exposición es significativa. Un estudio académico de 2025 identificó más de 11.000 LLM sin censura alojados en Hugging Face, con las variantes principales descargadas más de 19 millones de veces.[106] El modelo R1 de peso abierto de DeepSeek ilustra el límite máximo que puede alcanzar esta categoría: una evaluación de seguridad de Cisco descubrió que no logró bloquear ni una sola solicitud dañina de entre 50 intentos, y los investigadores de Enkrypt AI documentaron que generaba un blog detallado de reclutamiento terrorista en respuesta a una solicitud de prueba.[107,108]

Para fines de reclutamiento criminal, los modelos autoalojados ofrecen toda la capacidad generativa sin ninguna de las restricciones del lado del proveedor.

CRIMEN COMO SERVICIO (CAAS)

Crimen como servicio (CaaS) se refiere a la comercialización de capacidades delictivas, ya sea mediante suscripción o no, en la que se venden herramientas, conocimientos especializados y soporte técnico al usuario final. Aplicado a la IA, este modelo ha dado lugar a un mercado clandestino por niveles de herramientas delictivas diseñadas específicamente que empaquetan capacidades de IA para el fraude, la ingeniería social y la explotación.

La mayoría de estas herramientas se basan en el *jailbreaking*: técnicas que eluden las barreras de seguridad de los modelos de IA comerciales para producir contenidos que, de otro modo, rechazarían. Los métodos circulan libremente en foros clandestinos.[109,110] El Informe sobre amenazas de IA de 2025 de KELA documentó un aumento del 200% en las menciones de herramientas de IA maliciosas en foros de ciberdelincuencia monitorizados, junto con un incremento del 52% en los debates relacionados con el *jailbreaking*.[111]

104 Schultz, J. (2025). El uso indebido de los grandes modelos de lenguaje por parte de los ciberdelincuentes.
105 Schultz, J. (2025). El abuso de los modelos de lenguaje a gran escala por parte de los ciberdelincuentes.
106 Lin, Z., Li, Z., Liao, X. y Wang, X. (2025). Consiglieres en la sombra: Comprensión del uso de modelos de lenguaje grandes sin censura en los delitos cibernéticos. arXiv:2508.12622. https://doi.org/10.48550/arXiv.2508.12622
107 Swanson, E. (2025). DeepSeek R1: Análisis de seguridad y protección.
108 Pearl, M., Brock, J. y Kumar, A. (marzo de 2025). Profundizando en los peligros de DeepSeek. Centro de Estudios Estratégicos e Internacionales.
109 KELA. (2025). Informe sobre amenazas de IA de 2025: Cómo los ciberdelincuentes están utilizando la tecnología de IA como arma. KELA Cyber Threat Intelligence.
110 Simonovich, V. (2025). Investigación sobre amenazas de Cato CTRL: variantes de WormGPT impulsadas por Grok y Mixtral.
111 KELA. (2025). Informe sobre amenazas de IA de 2025: Cómo los ciberdelincuentes están utilizando la tecnología de IA como arma. KELA Cyber Threat Intelligence.

El valor de las herramientas CaaS reside en la comodidad, el anonimato y la eliminación de cualquier obstáculo en las tareas dañinas, lo que las hace especialmente relevantes como mecanismo para reducir las barreras de acceso para los actores menos cualificados. Las herramientas abarcan una amplia gama de precios. WormGPT 4, orientado al phishing y al compromiso de correo electrónico empresarial, está disponible a partir de aproximadamente 50 dólares al mes.[112] En el extremo más sofisticado, Xanthorox AI, identificada por primera vez en foros de la darknet en el primer trimestre de 2025, funciona como una plataforma multimodelo autoalojada con cinco módulos especializados que abarcan la generación de código, el contenido de ingeniería social, el análisis de imágenes, la interacción por voz y el reconocimiento web en tiempo real en más de 50 motores de búsqueda. A diferencia de sus predecesoras, no se basa en modelos base existentes ni en infraestructura de nube pública, lo que reduce su detección y trazabilidad.[113] Los investigadores de seguridad señalan, sin embargo, que las afirmaciones de marketing sobre herramientas de este tipo suelen exceder las capacidades verificadas.[114] En el extremo gratuito, herramientas como KawaiiGPT no requieren pago ni conocimientos técnicos, y pueden estar operativas en cuestión de minutos.[115]

Se han identificado más de 212 variantes de modelos maliciosos activos en plataformas clandestinas.[116]

112 Palo Alto Networks Unit 42. (25 de noviembre de 2025). WormGPT 4 y KawaiiGPT: los nuevos LLM oscuros impulsan la automatización de la ciberdelincuencia. SecurityWeek.

113 Mascellino, A. (7 de abril de 2025). La IA Xanthorox de Darknet ofrece herramientas personalizables para hackers.
114 NETSCOUT. (19 de diciembre de 2025). DDoS por encargo y el uso en evolución de la IA.
115 Palo Alto Networks Unit 42. (25 de noviembre de 2025). WormGPT 4 y KawaiiGPT: los nuevos LLM oscuros impulsan la automatización de la ciberdelincuencia. SecurityWeek.
116 Velasco, C., García Periche, J., Gómez Gómez, J. D., y Bueno Benedí, M. (2025). Inteligencia artificial y delincuencia organizada (informe EL PACCTO 2.0). Expertise France y la Fundación para la Internacionalización de las Administraciones Públicas. https://doi.org/10.5281/zenodo.16740421

SEGMENTACIÓN SOCIAL E INGENIERÍA SOCIAL

La IA ha ampliado la capacidad de reclutamiento de delincuentes de dos formas interrelacionadas: al permitir la identificación y selección precisa de personas vulnerables, y al facilitar la creación a gran escala de identidades y relaciones falsas convincentes.

SEGMENTACIÓN SOCIAL

La IA proporciona a los actores delictivos acceso a capacidades de segmentación de la audiencia que antes solo estaban al alcance de los anunciantes comerciales: la capacidad de identificar a individuos específicos en función de perfiles de vulnerabilidad y ofrecerles contenido dirigido a gran escala, en todas las plataformas e idiomas.

La arquitectura de la plataforma como infraestructura de segmentación

Los delincuentes aprovechan la infraestructura de las redes sociales para el reclutamiento. Las plataformas de redes sociales segmentan de forma natural a las audiencias, mostrando contenidos a los usuarios en función de su comportamiento de búsqueda, su historial de interacción y los datos de su perfil. Los delincuentes se aprovechan de esto para mostrar material de reclutamiento a las audiencias objetivo. Por ejemplo, se ha documentado que el feed «Para ti» de TikTok muestra contenidos de reclutamiento de cárteles a adolescentes y jóvenes adultos de comunidades de bajos ingresos que interactúan con contenidos relacionados con la narcocultura o con la búsqueda de empleo.[117(,)118] La segmentación de las redes sociales también sirve para conectar a los actores delictivos con posibles reclutas; por ejemplo, se ha identificado la función de «amigos» recomendados de Snapchat como un vector para el reclutamiento para el tráfico de drogas en la UE.[119] La propia plataforma también puede servir para segmentar a la población, ya que el contexto profesional de LinkedIn segmenta de forma inherente a la audiencia y confiere una falsa credibilidad a los anuncios de empleo de una manera que las plataformas genéricas no hacen.[120]

Perfilamiento de vulnerabilidad

La IA permite la segmentación por marcadores psicológicos y socioeconómicos específicos, no solo por categorías demográficas. Las herramientas de IA detectan indicadores de dificultades económicas, aislamiento social y vulnerabilidad emocional a través de la elección de palabras, el sentimiento y el comportamiento en línea.[121] Esto puede complementarse con grandes conjuntos de datos extraídos.[122]

Esto permite una segmentación precisa. Por ejemplo, la red criminal 764 buscaba sistemáticamente en X (antes Twitter) a usuarios que se identificaran públicamente como personas que se autolesionaban, dirigiéndose específicamente a chicas con depresión, trastornos alimentarios o ideas suicidas.[123] En una operación conjunta de Europol y las policías francesa y española en

117 Seminario sobre Violencia y Paz (2025), Reclutamiento criminal en TikTok: una investigación documenta más de 100 cuentas activas en México.
118 Tynan (2026), Cómo utilizan los cárteles mexicanos el contenido generado por IA.

119 Europol. (2025e). Robar, vender y repetir: cómo los ciberdelincuentes comercian y explotan tus datos — Evaluación de la amenaza de la delincuencia organizada en Internet (IOCTA) 2025. Oficina de Publicaciones de la Unión Europea.
120 Amerhauser y Goodwin (2026), Un mundo de engaños; OSCE OSR/CTHB (2026), La trata de personas en las operaciones de estafa cibernética.
121 Levesque, J. (2025). La amenaza del tráfico potenciada por la IA: Examen de la evolución tecnológica de las operaciones de trata de personas con herramientas de IA. Revista sobre la Trata de Personas. https://doi.org/10.1080/23322705.2025.2572911; Bender, E. M., Gebru, T., McMillan-Major, A., & Shmitchell, S. (2021). Sobre los peligros de los loros estocásticos: ¿pueden los modelos de lenguaje ser demasiado grandes? En Actas de la Conferencia ACM 2021 sobre Equidad, Responsabilidad y Transparencia (pp. 610–623). https://doi.org/10.1145/3442188.3445922
122 Levesque, J. (2025). La amenaza del tráfico potenciada por la IA: Examen de la evolución tecnológica de las operaciones de trata de personas con herramientas de IA.
123 Instituto para el Diálogo Estratégico. (2025). Redes de daño: un recurso de información centrado en las víctimas sobre la red de sextorsión 764.

TIPOLOGÍA	EDAD OBJETIVO	PRINCIPALES PLATAFORMAS DOCUMENTADAS
CSAM / OCSEA	0 – 17	Telegram, WeChat, Discord, Roblox
Redes 764 / NVE	10 – 17	Discord, Roblox, Telegram, X, Reddit
Tráfico de drogas (UE)	7 – adultos jóvenes	Instagram, Snapchat
VaaS (EU)	13+ incluidos adultos	TikTok, Instagram, Snapchat → Telegram, Signal, Wickr
Reclutamiento de cárteles (ALC)	11+ incluidos adultos	TikTok, Instagram, Facebook, plataformas de videojuegos (Free Fire, GTA V) → WhatsApp, Telegram
Explotación sexual	Adolescentes – adultos	Facebook, Instagram, plataformas de citas, sitios de anuncios clasificados
Centros de estafa	Adultos (principalmente)	Telegram, WhatsApp, LINE, Facebook, LinkedIn, Viber

Tabla 1. Tipologías de captación criminal por edad objetivo y principales plataformas documentadas

* Las plataformas enumeradas no constituyen una lista completa ni exhaustiva. Reflejan informes de fuentes abiertas y revelaciones de las fuerzas del orden procedentes de la tipología de este informe y tal y como se recogen únicamente en los estudios de caso. Debe suponerse la existencia de plataformas adicionales.

2022, una red de trata utilizó la IA para analizar los perfiles psicológicos y socioeconómicos de posibles víctimas en las redes sociales, segmentando a las mujeres en situaciones vulnerables antes de enviarles mensajes de captación personalizados.[124] Estos casos reflejan un nuevo tipo de lógica operativa que la IA hace escalable.

Dentro de estos perfiles de vulnerabilidad, la edad funciona como una variable operativa estructurante. Determina qué plataformas utiliza un objetivo, qué apelaciones psicológicas son efectivas, qué función delictiva puede desempeñar una persona reclutada y qué riesgo legal asume el reclutador. Las herramientas de IA se ajustan a estas distinciones: para los niños, a través de infraestructuras de captación y sistemas de recomendación de contenidos que explotan la dependencia y la capacidad limitada para dar su consentimiento; para los adolescentes, a través del contacto basado en plataformas, vectores de entrada entre pares y una interacción impulsada por la IA adaptada a su forma de comunicarse y establecer confianza; para los adultos, a través de identidades falsas, fraude laboral y coacción financiera. Se trata de diferentes modelos operativos aplicados a distintos perfiles de riesgo, con un patrón de selección basado en la edad coherente en tipologías delictivas por lo demás distintas, como muestra la Tabla 1.

Como ilustra la Tabla 1, los adolescentes aparecen como un grupo objetivo en múltiples tipologías distintas.

INGENIERÍA SOCIAL

Una vez identificado el objetivo, la IA permite a los delincuentes crear y mantener identidades y relaciones falsas a una escala y con un nivel de sofisticación que los métodos manuales no pueden igualar.

124 Velasco et al. (2025), Inteligencia artificial y delincuencia organizada.

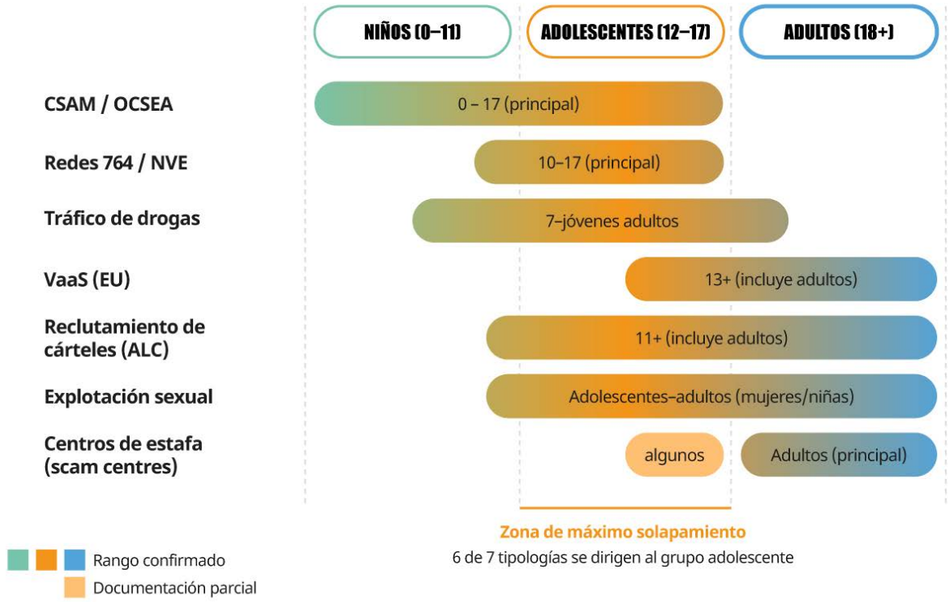


Figura 6. Espectro de edades objetivo

La figura 6 ilustra los rangos de edad a los que se dirigen principalmente las tipologías delictivas examinadas en este informe, así como la concentración del riesgo en el grupo de los adolescentes.

Relaciones inventadas

Las imágenes de perfil generadas por IA, los historiales profesionales falsos y las conversaciones impulsadas por modelos de lenguaje grande (LLM) permiten a un solo operador mantener múltiples identidades falsas simultáneamente en distintas plataformas.[125] Herramientas como SCRM Champion (véase la sección 3.1.1 – Centros de estafa) agrupan WhatsApp, LINE, Telegram, Facebook Messenger e Instagram en un único panel de control, con un asistente de IA que genera respuestas automáticas, realiza un seguimiento de los objetivos a través de un sistema de gestión de clientes potenciales y permite un traspaso fluido entre operadores, lo que hace que el proceso de generar una falsa confianza sea a la vez industrializado y difícil de detectar.[126] Los chatbots de IA pueden mantener estas interacciones de forma autónoma o con una supervisión humana mínima.

Falsa legitimidad

La IA permite a los reclutadores delictivos producir comunicaciones contextualmente adecuadas, gramaticalmente correctas y culturalmente adaptadas que eluden las señales que se ha enseñado a las víctimas a detectar. Las organizaciones delictivas pueden publicar anuncios de empleo elaborados con IA, respaldados por sitios web de empresas falsos, procesos de entrevista estructurados y documentos, creando una apariencia de legitimidad en varias capas que a los objetivos les resulta difícil

125 Levesque, J. (2025). La amenaza del tráfico potenciada por la IA: Examen de la evolución tecnológica de las operaciones de trata de personas con herramientas de IA.

126 Amerhauser y Goodwin (2026), Un mundo de engaños.

distinguir de un reclutamiento genuino.[127] La suplantación de identidad en LinkedIn mediante fotografías de perfil generadas por IA y cuentas pirateadas aprovecha la credibilidad institucional de la plataforma, lo que refuerza aún más esta imagen.[128]

RECLUTAMIENTO TOTALMENTE IMPULSADO POR IA

El reclutamiento totalmente impulsado por IA elimina por completo al operador humano del proceso de reclutamiento, lo que requiere que los agentes de IA lleven a cabo de forma autónoma la selección de objetivos, la interacción y el establecimiento de confianza con las víctimas potenciales.

Este modelo es incipiente, pero operativamente creíble. Las pruebas académicas indican que los agentes de IA son significativamente más eficaces que los humanos a la hora de generar confianza y cumplimiento en un entorno de reclutamiento,[129] , y los participantes no muestran ninguna sospecha durante las propias conversaciones.[130]

Los agentes LLM están disponibles sin interrupción y pueden gestionar miles de conversaciones simultáneas. Estos sistemas pueden operar a través de un proceso que se refuerza a sí mismo y que perfecciona continuamente los mensajes, la selección de canales y la selección de objetivos. El resultado es un sistema que se vuelve progresivamente más eficaz con el tiempo y cada vez más difícil de detectar para los agentes de seguridad.[131]

Este modelo operativo emergente sigue requiriendo actualmente un «moderador» humano que ajuste los guiones, identifique los perfiles demográficos de los objetivos e intervenga solo cuando sea necesario.[132] Esto

127 Amerhauser y Goodwin (2026), Un mundo de engaños; OSCE OSR/CTHB (2026), La trata en las operaciones de estafa cibernética.
128 Amerhauser y Goodwin (2026), Un mundo de engaños.
129 Gressel et al. (2025), Amor, mentiras y modelos de lenguaje. En un estudio ciego de una semana de duración, un agente LLM suscitó una confianza significativamente mayor entre los participantes (p = 0,007) y alcanzó una tasa de cumplimiento de tareas del 46 %, en comparación con el 18 % de los operadores humanos. Los operadores de modelos de IA señalaron el 0,0 % de las interacciones.
130 Gressel et al. (2025), Amor, mentiras y modelos de lenguaje.
131 Levesque (2025a), La amenaza de la trata potenciada por la IA.
132 Amerhauser y Goodwin (2026), Un mundo de engaños.

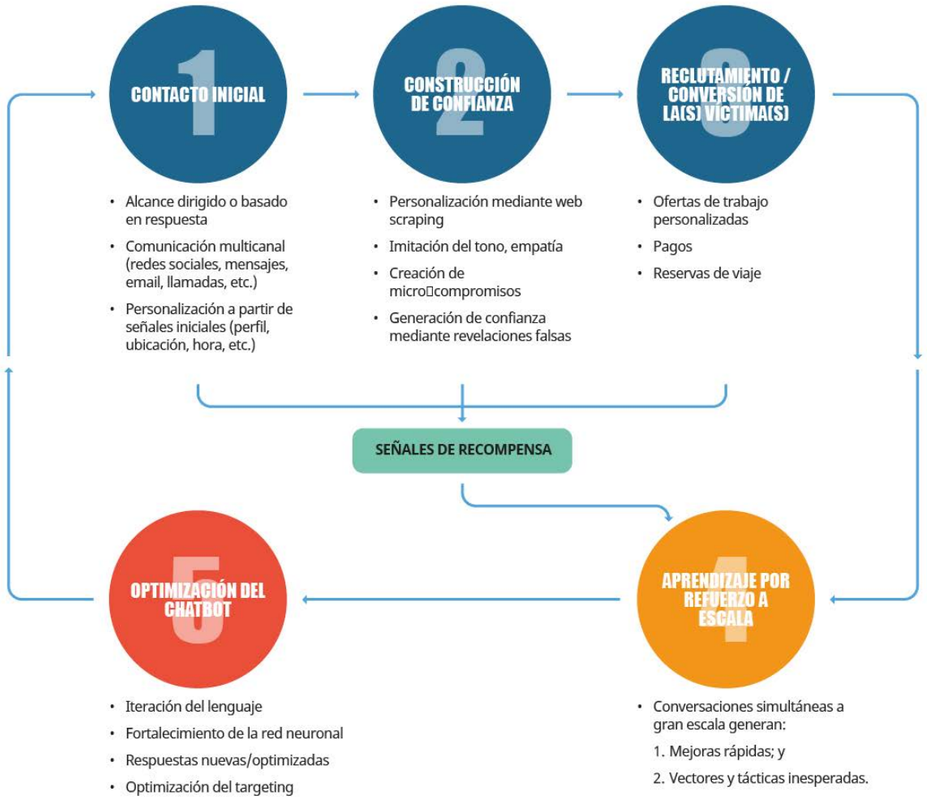


Figura 7. Canal de captación mediante chatbot con IA y bucle de aprendizaje por refuerzo.

**Fuente*: Levesque, J. (2025). La amenaza de la trata potenciada por la IA: análisis de la evolución tecnológica de las operaciones de trata de personas con herramientas de IA. *Journal of Human Trafficking*. https://doi.org/10.1080/23322705.2025.2572911

preserva el control humano a nivel estratégico, al tiempo que reduce drásticamente los costes operativos y la exposición. Un solo operador que implemente un sistema de este tipo puede alcanzar un alcance que, de otro modo, requeriría una plantilla humana numerosa y logísticamente compleja.[133]

La principal implicación es estructural. La mano de obra humana ha sido históricamente la limitación determinante en el reclutamiento. El reclutamiento totalmente automatizado elimina esa limitación, reduciendo la barrera de entrada para nuevos actores y permitiendo que las redes existentes se expandan sin un aumento proporcional del riesgo operativo.

133 Levesque, J. (2025). La amenaza del tráfico potenciada por la IA: Examen de la evolución tecnológica de las operaciones de trata de personas con herramientas de IA.

REDES CRIMINALES Y MODUS OPERANDI

Las redes criminales examinadas en esta sección abarcan distintos contextos operativos, huellas geográficas y modelos de explotación. Los siguientes estudios de caso documentan cómo funciona esa integración en la práctica, basándose en informes de fuentes abiertas, revelaciones de las fuerzas del orden e investigación primaria en las regiones de América Latina y el Caribe y la Unión Europea.

ESTUDIOS DE CASO

Los estudios de caso examinan el reclutamiento asistido por IA en diferentes tipologías delictivas. En conjunto, ilustran cómo se están aplicando las herramientas de IA en las diferentes etapas del proceso de reclutamiento.

CENTROS DE ESTAFA

Los centros de estafa contratan a trabajadores bajo falsos pretextos para llevar a cabo operaciones fraudulentas, entre las que se incluyen estafas financieras, estafas sentimentales, esquemas de inversión en criptomonedas, sextorsión y extorsión. Lo que distingue a los centros de estafa de la mayoría de las demás tipologías delictivas es su estructura de doble víctima. Explotan a los trabajadores víctimas de la trata como mano de obra forzada y, al mismo tiempo, utilizan a esos trabajadores para estafar a una población externa de víctimas financieras. Actualmente hay pruebas de que la IA potencia ambas dimensiones: cómo se recluta a los trabajadores en los recintos y cómo estos trabajadores seleccionan y manipulan a las víctimas una vez dentro.

Tecnologías utilizadas

Hay pruebas del uso de la IA en todas las fases del reclutamiento. Los conjuntos de herramientas basadas en la IA se comercializan, se integran en paquetes tecnológicos y se diseñan deliberadamente para ocultar la identidad de los operadores. Las tecnologías destinadas a defraudar a las víctimas de las estafas también sirven como mecanismos de reclutamiento en sí mismas.[134,135]

La suite 007TG, documentada por C4ADS a través de entrevistas directas con estafadores que operan en Myanmar y Filipinas, ofrece dos servicios principales: aumentar el alcance ampliando la presencia en línea del operador y ocultar los medios para hacerlo. Sus herramientas permiten a los operadores gestionar cientos de cuentas de redes sociales simultáneamente, automatizar la divulgación en más de 100 idiomas, generar números de teléfono aleatorios para la verificación de cuentas y procesar pagos en criptomonedas, todo ello mientras evaden la detección de las plataformas.

La plataforma afiliada KT Smart Translation integra cuatro servicios de traducción basados en modelos de lenguaje grande (LLM) (Google Translate, DeepL, ChatGPT y Papago de Corea del Sur), lo que permite a los operadores reclutar personal de distintos grupos lingüísticos sin necesidad de contar con personal multilingüe. Al eliminar la IA las barreras lingüísticas, los operadores pueden reclutar trabajadores independientemente de su origen lingüístico, ampliando considerablemente la reserva de mano de obra disponible.[136]

Las herramientas específicas de la suite 007TG facilitan directamente la contratación:

- **WorkGram** permite a los operadores gestionar varias cuentas de Telegram simultáneamente, enviar mensajes masivos con un solo clic y realizar traducciones en directo asistidas por IA.[137]
- **SCRM Champion** agrupa WhatsApp, LINE, Telegram, Facebook Messenger e Instagram en un único panel de control, con un asistente basado en IA que genera respuestas automáticas, sugiere mensajes de seguimiento y realiza un seguimiento de los objetivos a través de un sistema de gestión de clientes potenciales, lo que permite un traspaso fluido de una víctima potencial de un operador a otro.[138]
- **Promo Picasso** ofrece envío masivo de mensajes con gestión centralizada de cuentas, asignando direcciones IP independientes a cada cuenta para reducir el riesgo de bloqueos de la plataforma.[139]
- **SMSBower** proporciona una infraestructura de tarjetas SIM virtuales y temporales para la creación y verificación de cuentas a gran escala.[140] Al implementar estas herramientas de forma combinada, un operador puede crear una presencia a gran escala en las redes sociales, recopilar perfiles detallados de los objetivos a partir de datos extraídos, llevar a cabo una divulgación automatizada y personalizada, y gestionar todas las comunicaciones desde detrás de capas de números virtuales e IP rotativas, sin que sea posible identificarlo.[141]

Existen pruebas de que se utiliza la IA en la fase de entrevistas del proceso de selección.[142] La tecnología de «cambio de rostros» de Haotian permite a los operadores superponer el rostro de otra persona en transmisiones de vídeo en directo, presentando un avatar generado por IA como reclutador o empleador.[143,144]

134 Di Girolamo (2026), Engañoso por diseño.
135 OSCE OSR/CTHB (2026), La trata en las operaciones de estafa cibernética.
136 Amerhauser y Goodwin (2026), Un mundo de engaños.
137 Di Girolamo (2026), Engañoso por diseño.
138 Di Girolamo (2026), Engañoso por diseño.
139 Di Girolamo (2026), Engañoso por diseño.
140 Di Girolamo (2026), «Engañoso por diseño».
141 Di Girolamo (2026), Engañoso por diseño.
142 Según los informes de los supervivientes y los expertos entrevistados, mayo de 2026.
143 OSCE OSR/CTHB (2026), La trata de personas en las operaciones de estafa cibernética.
144 Cox, J. (7 de mayo de 2026). «HELLO BOSS»: Dentro del software chino de deepfakes en tiempo real que impulsa estafas en todo el mundo. 404 Media.

Figura 8. Anuncio de Haotian en el que se muestra la tecnología «faceswap».

Fuente: 404 media

A la izquierda se ve a la persona real. A la derecha, un rostro superpuesto. El vídeo muestra a la modelo moviendo rápidamente la mano delante de su rostro, tocándose la cara y cambiando la iluminación, todo ello sin que el rostro superpuesto presente fallos.

Los centros de estafa pueden ampliar el número de caras deepfake mediante tres mecanismos:

1. Se pueden extraer rostros generados por IA de los perfiles de las redes sociales;
2. Es posible que se escaneen y suban las caras de los trabajadores de los centros de estafas;[145]
3. Las personas solicitan trabajar en centros de estafa para ser modelos de IA.[146]

En combinación con la clonación de voz mediante IA, esto permite crear un entrevistador totalmente ficticio.[147] Cabe señalar que la captación de modelos de IA a través de canales de Telegram es una vía de reclutamiento cada vez más frecuente para las propias operaciones de los centros de estafa,[148] y están surgiendo informes sobre trabajo forzoso y otras formas de explotación y abuso (incluido el abuso sexual) una vez allí.[149,150] Entre los solicitantes de «modelos de IA» se encontraban personas de países

145 Señalado en una entrevista con un superviviente de un centro de estafas. Mayo de 2026.
146 Basado en informes de supervivientes, tal y como se detalla en una entrevista realizada en mayo de 2026.

147 INTERPOL. (2024). Evaluación de fraude financiero de INTERPOL: Una amenaza global impulsada por la tecnología.
Remote Work Europe. (s. f.). Guía de estafas de trabajo remoto.
148 Malwarebytes. (2026, 23 de marzo). Los complejos de estafa contratan "modelos de IA" para cerrar acuerdos en videollamadas con deepfakes.Malwarebytes Labs.
149 Entrevista con un superviviente de un centro de estafas. Mayo de 2026.
150 Burgess, M. (2026, 16 de marzo). Las modelos se postulan para ser la cara de estafas con IA. WIRED.

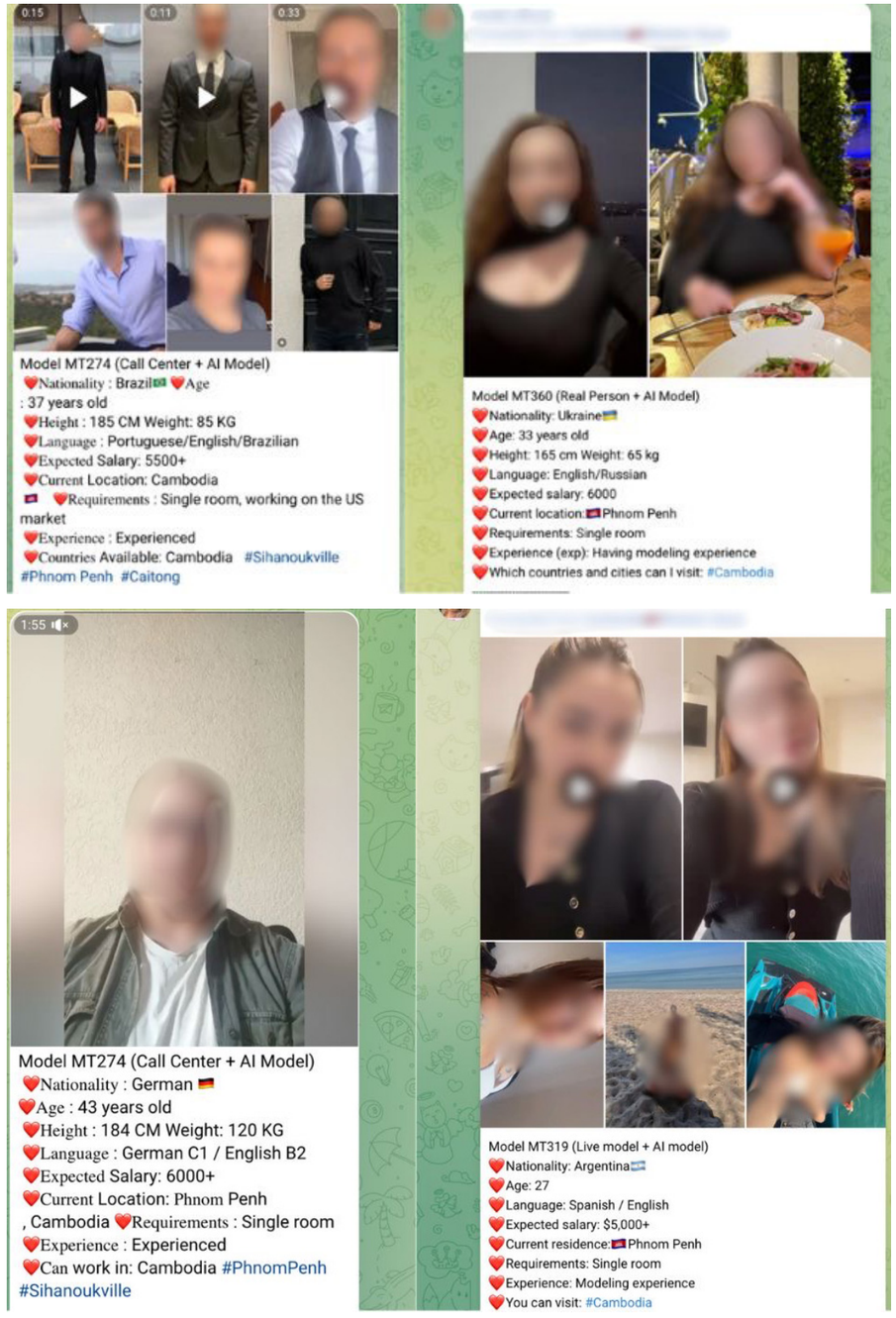


Figura 9. Capturas de pantalla de un grupo de Telegram con perfiles de personas que solicitan convertirse en «modelos de IA» para centros de estafas.

Las publicaciones seleccionadas muestran a personas de Brasil, Ucrania, Alemania y Argentina. Sus publicaciones incluyen los salarios que esperan recibir por sus servicios como modelos de IA, que oscilan entre los 5000 y los más de 6000 dólares estadounidenses.

de América Latina y el Caribe y Europa, como Argentina, Bielorrusia, Brasil, Colombia, Alemania, Rusia, España, Ucrania y el Reino Unido, así como personas de Jordania, Sudáfrica, Turquía, Turkmenistán, Uzbekistán y varios países asiáticos.

En cuanto a la captación, los centros de estafa utilizan anuncios de empleo creados por IA en plataformas generalistas, como Facebook, Telegram, LinkedIn, WhatsApp y Viber.[151,152] Los anuncios elaborados por IA utilizan títulos de puestos eufemísticos, como gestor de servicios, especialista en TI o agente de atención al cliente. Las fotos de perfil generadas por IA y la captación impulsada por IA permiten enviar mensajes de captación personalizados según el contexto a través de perfiles de aspecto realista. También se utilizan perfiles pirateados de profesionales reales para añadir credibilidad institucional a la captación.[153]

Se utilizan ampliamente los grupos de reclutamiento de Telegram, en los que se hace creer a los trabajadores que están accediendo a un empleo legítimo en telemarketing o tecnología.[154] A continuación se muestra un ejemplo de publicación.

También hay indicios cada vez más evidentes de que se está actuando de forma selectiva contra personas concretas. Los delincuentes están utilizando modelos de inteligencia artificial para rastrear las redes sociales en busca de personas que muestren indicios de dificultades económicas o problemas laborales, con el fin de aprovechar estas vulnerabilidades para llevar a cabo procesos de contratación fraudulentos.[155]

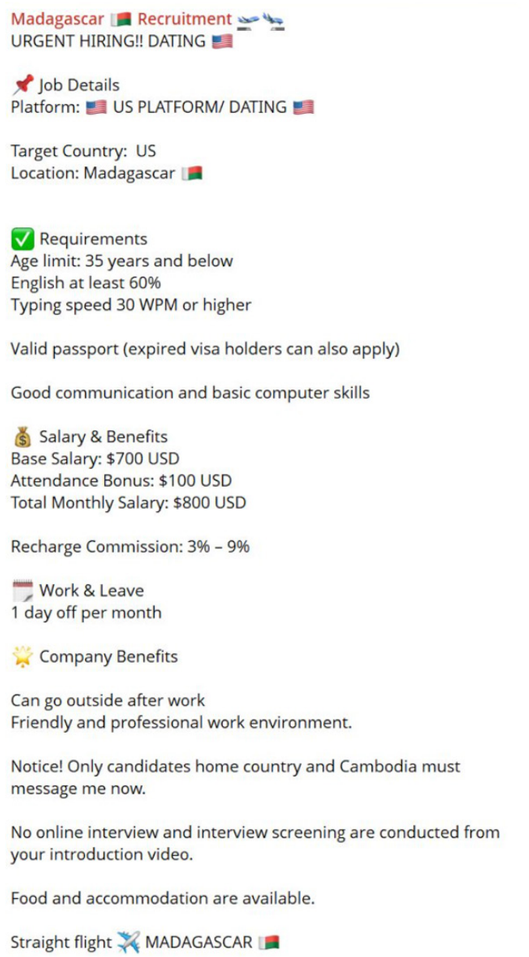


Figura 10. Oferta de empleo creada por IA publicada en una cuenta de Telegram vinculada a un centro de estafas

Crédito: Erin West. Publicación en LinkedIn, mayo de 2026. El anuncio de contratación es representativo del tipo de publicaciones que se ven. Está escrito por IA, como indica una puntuación de GPTZero del 100 % de texto generado por IA. Especifica un salario base de 800 USD al mes con comisiones por rendimiento del 3-9 %, requisitos de pasaporte, e incluye la frase: «No se realizan entrevistas en línea ni se lleva a cabo una selección a partir de tu vídeo de presentación». Se ofrecen manutención, alojamiento y un vuelo directo. La expresión «comisión de recarga» es una referencia directa a las cuotas de rendimiento fraudulentas.

151 OSCE OSR/CTHB (2026), Trata de personas en operaciones de estafa cibernética.
152 Fortune. (20 de noviembre de 2025). Centros de estafas del sudeste asiático, IA y trata de personas.
153 Entrevistas a expertos realizadas por el autor (mayo de 2026).
154 Amerhauser y Goodwin (2026), Un mundo de engaños.
155 Oficina del Alto Comisionado de las Naciones Unidas para los Derechos Humanos. (2026). «Un problema complejo»: en busca de soluciones basadas en los derechos humanos para la trata de personas destinada a operaciones de estafa cibernética en el sudeste asiático. Naciones Unidas.

Una vez que los trabajadores se encuentran dentro de los recintos, un conjunto independiente de herramientas de IA controla a la plantilla. CloudSeven (también conocida como Cloud007), una plataforma afiliada a 007TG, permite a los supervisores supervisar de forma remota la actividad en los ordenadores de los empleados, generar capturas de pantalla aleatorias y elaborar informes de datos de comportamiento generados por IA. El análisis del código fuente realizado por C4ADS reveló que CloudSeven integra más de 20 modelos de IA, entre los que se incluyen Claude, GPT, Gemini, DeepSeek y Llama.[156] SCRM Champion complementa esto con un análisis de sentimientos que rastrea los cambios emocionales en las respuestas de los objetivos, proporcionando a los supervisores visibilidad en tiempo real sobre si los estafadores están cumpliendo con sus cuotas.[157]

Modus operandi

El proceso de reclutamiento puede variar. En los procesos de reclutamiento más informales, las ofertas de trabajo fraudulentas se envían a través de plataformas de mensajería. Una vez que una persona responde, el proceso avanza rápidamente y las víctimas reciben ofertas de trabajo y se espera que viajen en cuestión de días.[158] También existen vías de reclutamiento más estructuradas, en las que, tras el anuncio inicial, un reclutador se pone en contacto directamente a través de WhatsApp, Telegram, mensajes directos de Instagram o llamadas telefónicas. El proceso de selección posterior imita el reclutamiento profesional y las entrevistas de trabajo legítimas: videoconferencias estructuradas con un panel de dos o tres entrevistadores, un entrevistador principal que habla inglés y uno o dos observadores silenciosos. Se evaluaba a los candidatos en cuanto a velocidad de mecanografía, situaciones de atención al cliente y dominio del idioma. Algunos se sometían a dos rondas de entrevistas, lo que reforzaba considerablemente la percepción de legitimidad.[159]

EXPLOTACIÓN Y ABUSO SEXUAL INFANTIL EN LÍNEA

La IA generativa ha alterado de forma fundamental el panorama del reclutamiento y la captación para la OCSEA. Lo que antes requería acceso directo a las víctimas, habilidades técnicas o un esfuerzo manual sostenido, ahora puede automatizarse, personalizarse y desplegarse a gran escala. La IA es un elemento central de cómo se inicia, intensifica y mantiene el reclutamiento.

Tecnologías utilizadas

Las principales herramientas documentadas en el reclutamiento para la OCSEA se dividen en cuatro categorías.

- *Generación y manipulación de imágenes.* Los modelos de difusión de código abierto, incluidos Stable Diffusion y sus derivados, se utilizan para generar material fotorrealista de abuso sexual infantil (CSAM) a partir de indicaciones de texto y para convertir fotografías inocuas de niños en material sexualmente explícito. Las aplicaciones específicas de «desnudez», que eliminan la ropa de las fotografías mediante IA, están ampliamente disponibles en las principales tiendas de aplicaciones y no requieren ninguna configuración técnica. Estas herramientas se utilizan tanto para producir material de explotación como para generar imágenes fabricadas de víctimas específicas con el fin de utilizarlas como medio de chantaje sexual.[160,161]
- *Chatbots de IA y simulación de personalidades.* Plataformas como Character.AI albergan

156 Di Girolamo (2026), Engañoso por diseño.
157 Di Girolamo (2026), Engañoso por diseño.
158 Según lo detallado por un experto de un centro regional contra el fraude, basado en entrevistas con supervivientes. Mayo de 2026.
159 Entrevista con un experto de un centro regional contra el fraude, mayo de 2026.
160 Fundación Internet Watch. (2025b). Imágenes generadas por IA. En Informe anual de datos y análisis 2025.
161 Aoukar, Y., Péron, C., Maddox, L., Apple, L. y Centro de Inteligencia Artificial y Robótica del UNICRI. (2024). IA generativa: una nueva amenaza para la explotación y el abuso sexual infantil en línea. Fundación Bracket y Value for Good.

chatbots que simulan relaciones entre iguales con niños. Cuando se accede a ellos a través de cuentas de registradas a nombre de menores, se ha documentado que estos bots se involucran en una escalada de contenido sexualmente explícito, instruyen a los niños para que oculten las comunicaciones a sus padres y redirigen a los usuarios a canales no supervisados cuando intervienen los filtros de contenido.[162] También se han identificado plataformas creadas expresamente para albergar «personajes» de chatbot diseñados específicamente para simular escenarios sexuales con avatares infantiles, lo que representa un cambio que va más allá de las herramientas convencionales modificadas hacia una infraestructura construida para la explotación.[163]

- *Texto generado por IA.* Los modelos de lenguaje grande (LLM), incluidas variantes de código abierto sin moderar, se utilizan para generar guiones de captación, guías operativas de sextorsión y manuales paso a paso para el abuso sexual infantil. Estos materiales circulan en foros de la dark web y canales encriptados como técnicas reutilizables.[164,165]

- *Clonación de voz.* Las herramientas de síntesis de voz en tiempo real, ahora disponibles como aplicaciones de consumo, permiten a los agresores adultos hacerse pasar por niños en interacciones de audio en directo en plataformas de juegos como Roblox y Discord.[166]

Modus operandi

Imágenes deepfake generadas por IA de un niño específico, producidas a partir de fotografías extraídas de las redes sociales, que se presentan a la víctima como prueba de que ya existe material comprometedor. A continuación, se coacciona a la víctima para que produzca contenido sexual real bajo la amenaza de su distribución. Este modelo elimina la necesidad de obtener primero material genuino y acorta el tiempo entre el contacto inicial y el control coercitivo.[167] Los guiones de captación generados por IA se comparten dentro de las comunidades de agresores y son utilizados por miembros sin habilidades independientes de manipulación psicológica, lo que distribuye de manera efectiva técnicas sofisticadas a través de la red. En las plataformas de videojuegos, la clonación de voz en tiempo real permite a los depredadores presentarse como niños de la misma edad durante la fase inicial de creación de confianza, un vector que las herramientas de seguridad existentes de las plataformas no fueron diseñadas para detectar.[168]

En 2025, el Centro Nacional para Niños Desaparecidos y Explotados (NCMEC) recibió más de 1,5 millones de denuncias a través de CyberTipline que indicaban una relación entre la IA generativa y la explotación sexual infantil, incluyendo más de 3.000 casos que implicaban específicamente captación o abuso a través del chat. Las denuncias de explotación infantil relacionadas con la IA generativa presentadas a la Oficina de Investigaciones de Seguridad Nacional de EE. UU. se dispararon un 600 % en la primera mitad de 2025 en comparación con el año anterior. La IWF identificó 3.443 vídeos de material de abuso sexual infantil (CSAM) generados por IA en 2025, frente a los 13 de 2024.[169,170,171]

162 ParentsTogether Action y Heat Initiative. (2025). Cariño, por favor, vuelve pronto: Explotación sexual, manipulación y violencia en las cuentas infantiles de Character AI.
163 Fundación Internet Watch. (22 de septiembre de 2025c). Los chatbots de IA y el abuso sexual infantil: una llamada de atención para adoptar medidas de protección urgentes [Entrada de blog].
164 Fundación Internet Watch. (2025a). Extorsión con coacción sexual. En Informe anual de datos y análisis 2025.
165 Aoukar et al. (2024), IA generativa: una nueva amenaza para la explotación sexual infantil en línea.
166 Foro Económico Mundial. (2026). Los deepfakes en tiempo real están reescribiendo las reglas de la seguridad infantil.

167 Oficina Federal de Investigación. (2023). Actores maliciosos que manipulan fotos y vídeos para crear contenido explícito y esquemas de sextorsión [Anuncio de servicio público].
168 FEM (2026), Los deepfakes en tiempo real están reescribiendo las reglas de la seguridad infantil.
169 Fundación Internet Watch. (2026). Informe sobre el material de abuso sexual infantil generado por IA 2026: Daño sin límites.
170 Collier, K. (febrero de 2026). La crisis de la explotación infantil mediante IA ya está aquí. NBC News.
171 Henderson Vaughan, E. (31 de marzo de 2026). El trabajo nunca se detiene: un primer vistazo a los datos de 2025 del NCMEC [Entrada de blog].

La red Com y 764

The Com es un ecosistema transnacional y difuso de redes criminales en línea dedicadas a la explotación sexual infantil, la sextorsión, la ciberdelincuencia y la violencia motivada ideológicamente. Su subgrupo más documentado desde el punto de vista operativo, 764, se fundó en 2021 y se ha fragmentado en docenas de ramificaciones, entre las que se incluyen CVLT, 8884, Harm Nation, Leak Society, Kaskar, H3ll y Court, al tiempo que mantiene tácticas y canales de reclutamiento comunes. La red está moldeada por la Orden de los Nueve Ángulos (O9A), una organización británica neonazi, satanista y aceleracionista que promueve la violencia, la explotación sexual infantil y el colapso social como fines ideológicos. A finales de 2025, el FBI tenía más de 350 investigaciones activas en sus 55 oficinas locales, con detenciones en al menos 23 países.[172,173,174,175]

La red Com no se dedica exclusivamente a la explotación infantil. Las capacidades técnicas relacionadas con la ciberdelincuencia —como el intercambio de tarjetas SIM, la apropiación de direcciones IP, el doxing y el secuestro de cuentas— se integran directamente en las operaciones de la OCSEA. A medida que las herramientas de IA se van incorporando a las prácticas habituales de la ciberdelincuencia, los subgrupos de Com dedicados a la explotación infantil pueden acceder a ellas sin necesidad de desarrollar soluciones técnicas propias.[176,177]

172 Departamento de Justicia de EE. UU. (abril de 2025b). Detenidos y acusados los líderes de «764», un grupo global de sextorsión infantil [Comunicado de prensa].
173 Departamento de Justicia de EE. UU. (diciembre de 2025a). El líder del grupo extremista «764» se declara culpable de los cargos de RICO y explotación infantil [Comunicado de prensa].
174 Fox News. (2025). El FBI investiga a más de 350 personas vinculadas a la violenta red online «764».
175 Noyes, D. (2025). La amenaza «764»: el FBI afirma que hay depredadores que acechan a niños en los videojuegos y las redes sociales.
176 Kapko, M. (2025). Detienen y imputan a los líderes de 764, un grupo internacional dedicado a la sextorsión infantil.
177 CyberScoop (2024), Las autoridades federales investigan a 764.

Perfil de las víctimas

Las redes Com/764 se dirigen deliberadamente a menores de entre 8 y 17 años con indicadores específicos de vulnerabilidad: niñas con antecedentes documentados de depresión, autolesiones, trastornos alimentarios o ideas suicidas; jóvenes LGBTQ+ percibidos como socialmente aislados; niños en Roblox, Minecraft y Fortnite, donde las bases de usuarios jóvenes y la supervisión parental limitada crean entornos de captación accesibles; y usuarios de comunidades de autolesiones y trastornos alimentarios en X, Reddit y Discord que señalan públicamente su vulnerabilidad a través de hashtags y biografías de perfil.[178,179] Un estudio etnográfico del ISD sobre 764 canales de redes sociales reveló que uno de los líderes de 764 afirmaba haber encontrado a la mitad de sus víctimas en X buscando a usuarios que se identificaran explícitamente como personas que se autolesionaban.

Tecnologías utilizadas

El proceso de reclutamiento de Com/764 integra herramientas de IA en tres momentos distintos.

1. *Contacto inicial.* Las imágenes de perfil generadas por IA proporcionan identidades sintéticas de compañeros que superan la inspección visual. Las herramientas de clonación de voz en tiempo real, ahora aplicaciones de consumo general, permiten a los agresores adultos hacerse pasar por niños en interacciones de audio en directo en Roblox y Discord, las principales plataformas de reclutamiento de la red. Esto permite a los depredadores superar la verificación informal de confianza durante las primeras interacciones de captación antes de pasar al contacto por mensaje de texto.[180]

178 Real Policía Montada de Canadá. (2024). Grupos violentos en línea que explotan a niños y jóvenes..
179 Instituto para el Diálogo Estratégico (2025), Redes de daño.
180 FEM (2026), Los deepfakes en tiempo real están reescribiendo las reglas de la seguridad infantil.

Figura 11. Integración de la IA en el proceso de reclutamiento de Com/764

2. *Escalada y coacción.* Las imágenes deepfake generadas por IA de una víctima específica, producidas a partir de fotografías inocuas de redes sociales, se presentan como prueba de que ya existe material comprometedor. A continuación, se coacciona a la víctima para que produzca contenido sexual real bajo la amenaza de su distribución. El FBI confirmó que este modelo es habitual en las estafas de sextorsión dirigidas a menores.[181] Un estudio de UNICEF, ECPAT e INTERPOL realizado en 11 países reveló que al menos 1,2 millones de niños revelaron que sus imágenes habían sido manipuladas para crear deepfakes sexualmente explícitos durante el último año.[182] Las aplicaciones específicas de «nudificación» están ampliamente documentadas como mecanismo de coacción.[183]

3. *Retención y expansión.* El texto generado por IA se integra en la cultura establecida de la red de hacer circular guías operativas escritas. Las guías de sextorsión redactadas por miembros de 764 detallan enfoques psicológicos adaptados a perfiles específicos de víctimas, incluyendo guiones de cumplidos calibrados para comunidades con trastornos alimentarios y comunidades de autolesión.[184] El FBI documentó que miembros de 764 crearon un canal falso de Telegram para la prevención del suicidio con el fin de atraer a menores de edad con tendencias suicidas y recopilar información personal para su divulgación y extorsión.[185]

Implicaciones estratégicas

El caso Com/764 demuestra que las herramientas de IA se están incorporando a procesos de reclutamiento que ya eran operativamente eficaces, amplificando la escala, la velocidad y la verosimilitud en lugar de sustituir la explotación dirigida por humanos. Tres características estructurales distinguen la integración de la IA en esta red.

En primer lugar, la convergencia de las técnicas de ciberdelincuencia y la explotación infantil dentro de The Com significa que los avances en ingeniería social asistida por IA y suplantación de identidad, desarrollados para la ciberdelincuencia con motivaciones económicas, están a disposición de los subgrupos centrados en la OCSEA sin necesidad de desarrollo técnico adicional.

En segundo lugar, la cultura establecida de la red de producir y distribuir guías operativas escritas crea una vía natural de adopción para el contenido generado por IA: los guiones de captación, las guías de construcción de identidades y las plantillas de sextorsión producidas por IA pueden ser compartidas y desplegadas en toda la red por miembros sin capacidad técnica independiente.

En tercer lugar, la clonación de voz en tiempo real mediante IA en las plataformas de juego representa una brecha estructural en la detección que la infraestructura de seguridad actual de las plataformas no fue diseñada para abordar y que no ha recibido ninguna respuesta regulatoria específica.

CARTELES

Se ha vinculado a cinco importantes organizaciones mexicanas con el reclutamiento mediado digitalmente, con pruebas emergentes de la producción de contenido asistida por IA, en particular en el caso del Cártel de Jalisco Nueva Generación y el Cartel del Noreste (CDN).[186,187,188] Una investigación del *Seminario sobre Violencia y Paz de El Colegio de México* documentó que alrededor del 55 % de las cuentas de TikTok vinculadas a los cárteles estaban asociadas al Cártel de Jalisco Nueva Generación, y que el reclutamiento era la función principal en casi la mitad de los casos.[189] A pesar de ello, el Cartel del Noreste ha sido identificado como el usuario más activo de la IA para generar contenido en plataformas de redes sociales abiertas, principalmente TikTok e Instagram.[190,191,192]

Contenido generado por IA

Los cárteles utilizan la IA para producir y distribuir contenido de reclutamiento en forma de imágenes, vídeos y música generados por IA en las plataformas de redes sociales. El contenido se integra en la cultura musical del *narcocorrido,* las imágenes de estilo de vida y la iconografía popular, posicionando el reclutamiento criminal como un empleo o una identidad subcultural en lugar de una incitación a la delincuencia

Tecnologías utilizadas

El contenido de reclutamiento de los cárteles se produce utilizando herramientas de IA generativa disponibles en el mercado, capaces de crear imágenes, vídeos cortos y música a gran escala. Se trata de las mismas plataformas de doble uso que están disponibles públicamente, reutilizadas para producir contenido persuasivo a un volumen y una velocidad que los métodos manuales no pueden igualar.[193,194]

181 FBI (2023), Actores maliciosos que manipulan fotos y vídeos.
182 UNICEF. (febrero de 2026). «El abuso mediante deepfakes es abuso», advierte UNICEF.
183 Centro Canadiense para la Protección de la Infancia. (Octubre de 2024). La policía y la agencia de protección infantil dicen que los padres deben conocer los deepfakes sexuales generados por IA [Comunicado de prensa].
184 Instituto para el Diálogo Estratégico (2025), Redes de daño.
185 CyberScoop (2024), Las autoridades federales investigan el 764.
186 Seminario sobre Violencia y Paz (2025), Reclutamiento criminal en TikTok: una investigación documenta más de 100 cuentas activas en México.
187 Bledsoe (2025), El papel de las redes sociales en el reclutamiento de los cárteles.
188 Tynan (2026), Cómo utilizan los cárteles mexicanos el contenido generado por IA.
189 Seminario sobre Violencia y Paz (2025), Reclutamiento criminal en TikTok: una investigación documenta más de 100 cuentas activas en México.
190 Bledsoe (2025), El papel de las redes sociales en el reclutamiento de los cárteles.
191 Seminario sobre Violencia y Paz (2025), Reclutamiento criminal en TikTok: una investigación documenta más de 100 cuentas activas en México.
192 Tynan (2026), Cómo utilizan los cárteles mexicanos el contenido generado por IA.
193 Aguilar Antonio, J. M. (2025). Uso de la inteligencia artificial por parte de redes criminales de alto riesgo. Expertise France. https://doi.org/10.5281/zenodo.16750778
194 Tynan (2026), Cómo utilizan los cárteles mexicanos el contenido generado por IA.

El uso de contenido generado por IA cumple tres funciones principales:

- **Alcance.** Las herramientas de IA han transformado el volumen y la variedad de contenidos que los cárteles pueden producir. Las imágenes generadas por IA y las plantillas gráficas permiten a los reclutadores crear un gran número de publicaciones que pueden parecer profesionales, adaptadas al contexto local o personalizadas para diferentes públicos, lo que reduce el coste de producir contenidos persuasivos a gran escala.[195] Un mismo tema de reclutamiento puede replicarse en muchas variantes, cada una ajustada en tono, redacción o imágenes, y reutilizarse en TikTok, Instagram Reels y Facebook antes de dirigir a los usuarios interesados a grupos de WhatsApp o Telegram gestionados por coordinadores humanos.[196,197,198]

  Ese contenido generado por IA incluye vídeos de líderes de cárteles, figuras políticas y famosos diseñados para generar visualizaciones y comparticiones, y/o crear una sensación de legitimidad en torno a los esfuerzos de reclutamiento.[199] También incluye vídeos animados de personajes de videojuegos y películas infantiles, así como de conocidos personajes de películas de terror.[200] En TikTok, en particular, los clips cortos asistidos por IA están optimizados para el feed «Para ti», aprovechando la alta penetración entre los jóvenes y una moderación comparativamente débil para maximizar el alcance.[201] Estos clips se muestran a adolescentes y adultos jóvenes de comunidades de bajos ingresos que interactúan con contenidos relacionados con la narcocultura o la búsqueda de empleo, lo que aumenta la probabilidad de que los usuarios económica y socialmente vulnerables se vean expuestos a ofertas explotadoras.[202,203,204]

- **Compromiso.** Las imágenes, los vídeos, la música y los gráficos de marca generados por IA hacen que el contenido de reclutamiento parezca legítimo y culturalmente familiar. El uso de la IA en sí mismo se ha convertido, en algunos casos, en parte de la imagen de marca.[205] El contenido puede hacer que el reclutamiento parezca «aspiracional o económicamente atractivo», con publicaciones que se asemejan a anuncios de empleo, oportunidades de viaje o contenido sobre estilo de vida, en lugar de una incitación delictiva manifiesta.[206] Además, el uso de bots de IA para comentar o dar «me gusta» a las publicaciones (amplificación por bots) también puede influir en la percepción de legitimidad.[207]

- **Evasión.** El contenido generado por IA también funciona como una capa de evasión. Las publicaciones de reclutamiento se basan en combinaciones de señales como jerga afiliada a los cárteles, emojis codificados (por ejemplo, «😈», «👹» o «🐔»), marcas afiliadas estilizadas, capturas de pantalla de códigos QR, elecciones musicales y referencias geográficas, en lugar de nombres explícitos de cárteles.[208] Esta señalización simbólica permite que las publicaciones individuales parezcan inofensivas de forma aislada, sin dejar de ser reconocibles.[209,210] Si las hay, las ofertas de trabajo se colocan en las secciones de comentarios en lugar de en los pies de foto para reducir su visibilidad ante los moderadores de la plataforma.[211,212]

195 Tynan (2026), Cómo utilizan los cárteles mexicanos el contenido generado por IA.
196 Bledsoe (2025), El papel de las redes sociales en el reclutamiento de los cárteles.
197 Tynan (2026), Cómo utilizan los cárteles mexicanos el contenido generado por IA.
198 González, F. (25 de febrero de 2026). Cómo el cártel de la droga mexicano «CJNG» adoptó la IA, los drones y las redes sociales. WIRED.
199 Entrevista con un experto, abril de 2026.
200 Entrevista con un experto, abril de 2026. Cuentas de TikTok de reclutamiento para los cárteles que utilizan imágenes de terror, incluyendo hombres lobo con uniformes militares y personajes de películas como Chucky, que ofrecen detalles sobre oportunidades de trabajo con ofertas para cubrir los gastos de manutención.
201 Seminario sobre Violencia y Paz (2025), Reclutamiento criminal en TikTok: una investigación documenta más de 100 cuentas activas en México..
202 Tynan (2026), Cómo utilizan los cárteles mexicanos el contenido generado por IA.
203 Seminario sobre Violencia y Paz (2025), Reclutamiento criminal en TikTok: una investigación documenta más de 100 cuentas activas en México.
204 Tynan (2026), Cómo utilizan los cárteles mexicanos el contenido generado por IA.
205 Tynan (2026), Cómo utilizan los cárteles mexicanos el contenido generado por IA.
206 Tynan (2026), Cómo utilizan los cárteles mexicanos el contenido generado por IA.
207 Tynan (2026), Cómo utilizan los cárteles mexicanos el contenido generado por IA.

En un ejemplo notable, el Cártel de Jalisco Nueva Generación utilizó un vídeo generado por IA con personajes de dibujos animados antropomórficos, de aspecto similar al de una película infantil o un personaje de videojuego, en contraste con vehículos militares y personal armado, para anunciar el reclutamiento con ofertas de gastos de manutención cubiertos.[213] El uso de personajes no humanos aleja el contenido de la violencia explícita, ampliando su atractivo y haciendo que parezca más cercano al entretenimiento que a la incitación a la delincuencia.

Figura 12. Fotograma de un vídeo generado por IA vinculado al Cártel de Jalisco Nueva Generación, en el que se utilizan figuras antropomórficas para promover el reclutamiento. Fuente: Tynan (2026).

208 Tynan (2026), Cómo utilizan los cárteles mexicanos el contenido generado por IA.
209 Signa Lab ITESO (2024), TikTok y el reclutamiento para actividades criminales en México.
210 Tynan (2026), Cómo utilizan los cárteles mexicanos el contenido generado por IA.
211 Bledsoe (2025), El papel de las redes sociales en el reclutamiento de los cárteles.
212 Tynan (2026), Cómo utilizan los cárteles mexicanos el contenido generado por IA.
213 Tynan (2026), Cómo utilizan los cárteles mexicanos el contenido generado por IA.

Las campañas se han dirigido específicamente a hombres jóvenes de zonas económicamente marginadas, adolescentes interesados en contenidos sobre el estilo de vida de los cárteles, madres solteras a las que se les ofrece apoyo económico y personas en busca de trabajo.[214]

Modus operandi

Este contenido funciona como la parte superior de un embudo de reclutamiento. Una vez que el objetivo interactúa, a menudo se le traslada a canales privados o encriptados, normalmente WhatsApp o Telegram, donde un operador humano toma el relevo.[215] Existen algunos ejemplos documentados de cárteles que utilizan identidades falsas con fines de reclutamiento en Instagram.[216] Las pruebas iniciales también indican que se podrían estar utilizando chatbots con el fin de interactuar directamente, especialmente con adolescentes.[217]

CASO ESPECÍFICOS: ECUADOR

En Ecuador están empezando a surgir prácticas de reclutamiento asistidas por IA. Una encuesta realizada en 2025 por el Observatorio Ecuatoriano del Crimen Organizado reveló que al menos el 27% de los menores que forman parte de bandas fueron reclutados a través de las redes sociales, y la policía estima que hasta el 60% de los miembros de las bandas son adolescentes.[218] Las cuentas afiliadas a las pandillas utilizan imágenes generadas por IA que muestran representaciones simbólicas de animales correspondientes a la nomenclatura de las pandillas (*Los Tiguerones, Los Lobos, Las Águilas, Los Lagartos*), combinadas con

214 Bledsoe (2025), El papel de las redes sociales en el reclutamiento de los cárteles.
215 Tynan (2026), Cómo utilizan los cárteles mexicanos el contenido generado por IA.
216 News 4 Tucson KVOA. (s. f.). Una madre de Arizona con un mensaje para los padres después de que su hija fuera víctima de un «catfishing». KVOA.
217 Entrevista con un experto, abril de 2026.
218 InSight Crime. (6 de agosto de 2025). ¿Está América Latina preparada para mitigar el reclutamiento digital por parte de los grupos delictivos?

Figura 13. Imágenes de reclutamiento de pandillas generadas por IA procedentes de Ecuador.

Contenido generado por IA utilizado por pandillas ecuatorianas en TikTok que muestra una guerra simbólica entre animales (izquierda: águila/tigre) y composiciones relacionadas con pandillas (derecha: cabezas de león/águila con signos de pandillas y secuencias codificadas de emojis). Obsérvese los marcadores visibles de generación por IA. Cuentas afiliadas a Los Choneros, 2025-2026.

secuencias codificadas de emojis para indicar su afiliación y eludir la moderación de contenidos de la plataforma. El contenido observado incluye composiciones generadas por IA de animales asociados a las pandillas sobre figuras humanas que muestran signos de las pandillas, y personajes de películas de terror posicionados como personajes de reclutamiento. La función principal parece ser una representación simbólica y la evasión de la moderación, más que el engaño o la profesionalización.[219]

EXPLOTACIÓN SEXUAL

Los casos documentados de uso específico de la IA para reclutar víctimas con fines de explotación sexual son incipientes. Las pruebas disponibles apuntan a un patrón operativo emergente que merece una atención especial, sobre todo teniendo en cuenta el volumen del corredor de trata entre América Latina y el Caribe y la Unión Europea.[220,221]

La IA permite un enfoque de reclutamiento distribuido e independiente de la plataforma que amplía el alcance al tiempo que minimiza la exposición. A través de contenido generado por IA, identidades sintéticas y redes de cuentas falsas, los actores delictivos pueden mantener una presencia simultánea en plataformas de redes sociales, aplicaciones de citas y sitios de anuncios clasificados sin los costes de una gran red de reclutamiento humano. En cuanto a la selección de víctimas, las herramientas de IA rastrean las redes sociales y los perfiles de citas para generar perfiles socioeconómicos de las víctimas potenciales, identificando indicadores de vulnerabilidad como dificultades económicas,

219 Entrevista con un experto, mayo de 2026.

220 Velasco et al. (2025), Inteligencia artificial y delincuencia organizada.
221 Montes Delijorge (2026), «La IA está transformando el panorama de la explotación humana».

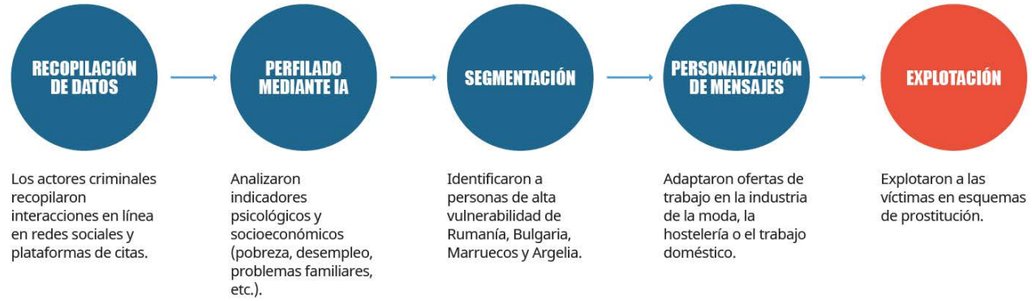


Figura 14. Cadena de explotación de la IA: de la elaboración de perfiles a la trata. Fuente: Levesque, J. (2025), basado en el ejemplo de caso descrito en Velasco et al. (2024, p. 39).

desempleo e inestabilidad familiar.[222,223] A continuación, se pueden generar y mantener a gran escala comunicaciones de reclutamiento personalizadas y culturalmente adecuadas, que imitan el lenguaje, el tono y el contexto de la víctima.[224,225] El contenido generado por IA también facilita la evasión: al modificar palabras clave, imágenes y textos publicitarios en tiempo real, los delincuentes eluden los sistemas de moderación de contenidos de las plataformas, mientras que los perfiles sintéticos aumentan la resistencia a la supresión de cuentas.[226]

En 2022, la Policía Nacional española y la Gendarmería Nacional francesa, en una operación conjunta con Europol, desmantelaron una sofisticada red de trata de personas que utilizaba inteligencia artificial para reclutar y explotar a mujeres jóvenes de Europa del Este y el norte de África. La red se especializaba en la explotación sexual y utilizaba algoritmos avanzados de IA para identificar, manipular y atraer a las víctimas a través de las redes sociales y las plataformas de citas en línea.

La red criminal utilizaba la IA para analizar las redes sociales y los perfiles de citas de las jóvenes, segmentando a aquellas que se encontraban en situaciones vulnerables. Los sistemas de IA filtraban estos datos a partir de las interacciones en línea de las víctimas, los comentarios en redes sociales y la actividad de búsqueda. Utilizando esta información, los delincuentes personalizaban los mensajes y ofrecían falsas oportunidades de trabajo en Europa Occidental, incluyendo empleos en la industria de la moda, la hostelería y el trabajo doméstico. Las víctimas procedían principalmente de Rumanía, Bulgaria, Marruecos y Argelia.

Una vez establecida la confianza, se convencía a las víctimas para que viajaran a España y Francia, donde se les confiscaban los documentos a su llegada y se les obligaba a trabajar en redes de prostitución.

La red también utilizaba la IA para eludir los sistemas de moderación de contenidos en las redes sociales y las plataformas publicitarias. Los algoritmos modificaban las palabras clave y el contenido de los anuncios para evitar ser detectados por los filtros automáticos, publicando en sitios web de anuncios clasificados y plataformas como Facebook e Instagram bajo la apariencia de ofertas de servicios aparentemente legítimas. Los anuncios se modificaban con frecuencia, adaptando el lenguaje a los contextos locales

222 Levesque, J. (2025). La amenaza del tráfico potenciada por la IA: Examen de la evolución tecnológica de las operaciones de trata de personas con herramientas de IA.
223 OSCE y RSO del Proceso de Bali (2024), Nuevas fronteras.
224 Levesque, J. (2025). La amenaza del tráfico potenciada por la IA: Examen de la evolución tecnológica de las operaciones de trata de personas con herramientas de IA.
225 Europol (2025a), El ADN cambiante de la delincuencia grave y organizada.
226 Velasco et al. (2025), Inteligencia artificial y delincuencia organizada.

y utilizando imágenes manipuladas para evadir los controles de contenido, lo que dificultaba enormemente a las autoridades el seguimiento de la actividad delictiva en tiempo real.[227]

VIOLENCIA COMO SERVICIO

El Grupo de Trabajo Operativo GRIMM de Europol describe la violencia como servicio como «accesible, escalable e impulsada por ecosistemas en línea que permiten el reclutamiento, la coordinación y la ejecución a través de las fronteras».[228]

Tecnologías utilizadas

El uso de la IA cumple dos funciones principales: ampliar el alcance de la delincuencia organizada y captar reclutas más allá de las barreras lingüísticas y culturales. También existe la posibilidad de que la IA se utilice para identificar y seleccionar a personas vulnerables, aunque las pruebas siguen siendo circunstanciales.

- **Alcance:** La música generada por IA y las imágenes de estilo de vida asociadas a redes delictivas se distribuyen en plataformas como TikTok, Instagram y Snapchat para normalizar la actividad delictiva y generar una primera interacción.[229] Las investigaciones observan «me gusta» y comentarios generados por bots de IA en publicaciones sobre música y estilos de vida delictivos. Los comentarios suelen glorificar los elementos relacionados con el crimen y las bandas. Este uso de la IA sirve para amplificar artificialmente este contenido, inflando su popularidad percibida y utilizándolo como mecanismo de ingeniería social.[230]

- **Interacción:** La IA se utiliza en ambos extremos de la relación entre reclutador y recluta. Los reclutadores y los instigadores utilizan herramientas de IA generativa para salvar la distancia lingüística y cultural entre ellos y los posibles reclutas en otros países.[231] La capacidad de comunicarse con personas de otros países constituye una parte clave de la infraestructura de la «violencia como servicio» y ayuda a la red a ocultar sus acciones. Los reclutados, entre los que se incluyen personas de tan solo trece años, utilizan la IA en sentido contrario: para traducir mensajes, generar texto y proyectar un nivel de capacidad o madurez que les haga parecer creíbles ante los manipuladores con los que, de otro modo, no podrían comunicarse de forma eficaz.[232] Este uso bidireccional reduce la distancia práctica entre actores que no comparten un idioma, un origen o una conexión previa comunes.

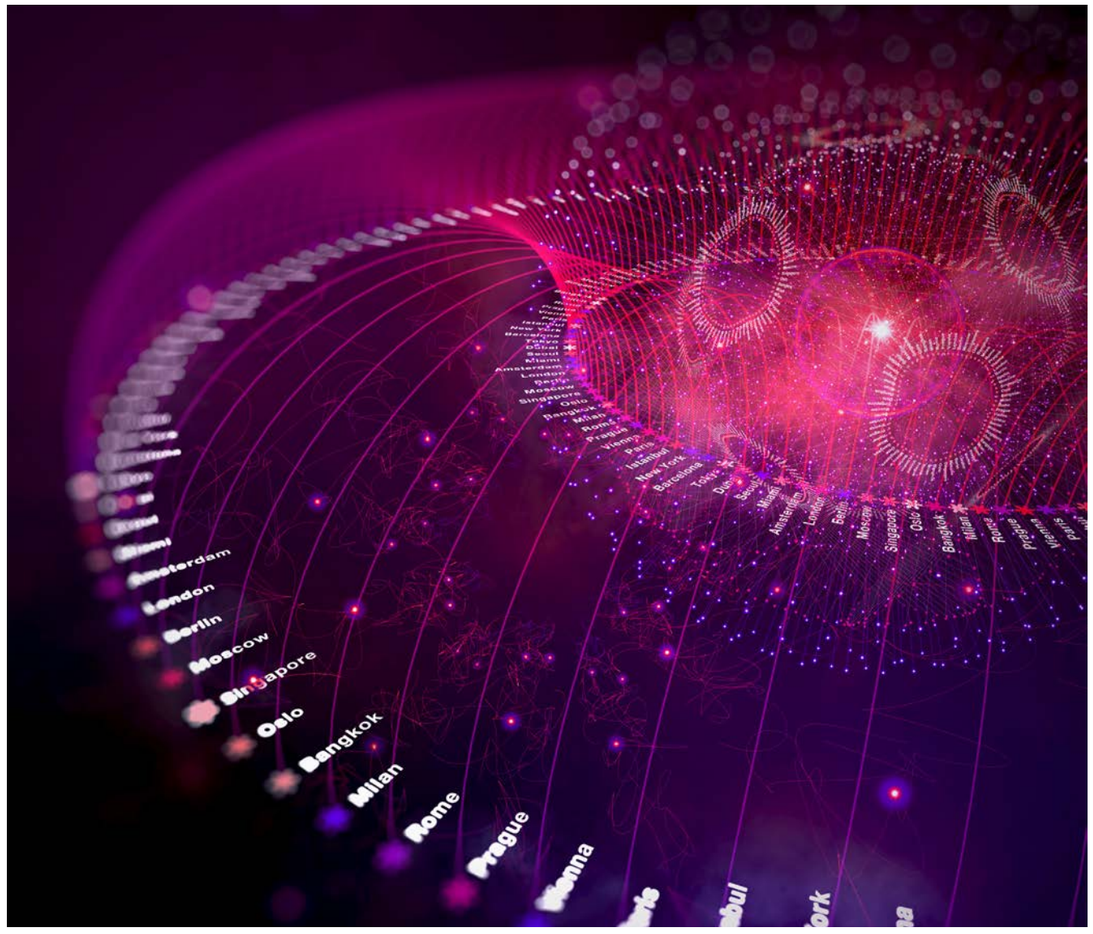
Estos grupos se dirigen sistemáticamente a personas vulnerables. Esto incluye comunidades online cerradas, como grupos de apoyo para jóvenes con problemas de salud mental. Estas comunidades son infiltradas y cultivadas y, en algunos casos, los grupos son vendidos a organizaciones criminales una vez que han acumulado un número de miembros suficientemente grande y vulnerable. Esto también incluye a adultos que comparten un perfil de vulnerabilidad común, como deuda financiera, dependencia previa de sustancias o aislamiento social agudo, y que se perciben a sí mismos como personas con pocas alternativas.[233] Esto constituye un ámbito preocupante en el que la IA podría implementarse para ampliar y aumentar la selección de adultos vulnerables.

Modus operandi

La arquitectura operativa comprende cuatro roles distintos: el instigador, que encarga el acto; el reclutador, que identifica y capta a los posibles autores a través de canales digitales; los facilitadores locales, que proporcionan la logística en la zona objetivo; y el autor, que ejecuta el acto.[234,235,236] La composición puede ser internacional y ningún participante necesita tener una relación previa con ningún otro. En un caso documentado, un equipo multinacional compuesto por ciudadanos irlandeses, ucranianos y neerlandeses se coordinó de forma remota para llevar a cabo un acto violento en un país vecino, con actores afiliados localmente que proporcionaron la logística, incluyendo alojamiento y armas.[237] El instigador y el autor no se conocían entre sí. El acto fue una transacción comprada. Esta estructura no se limita geográficamente a Suecia: el volumen de casos de OTF GRIMM abarca once países europeos participantes, incluidos Estados miembros de la UE y socios europeos vecinos, y redes vinculadas a Suecia han estado directamente implicadas en delitos violentos encargados en Dinamarca, Noruega y otros Estados vecinos.[238,239]

La IA no se limita a apoyar este modelo; lo hace viable a gran escala. Al reducir las barreras lingüísticas, permitir la publicidad masiva y facilitar la coordinación entre desconocidos de distintas jurisdicciones, la IA ayuda a las organizaciones criminales a externalizar la violencia de su propia plantilla, al tiempo que mantiene el control operativo. El producto se encarga de forma remota, mientras que la cadena de suministro funciona sobre una infraestructura digital. La IA facilita la industrialización del proceso.

227 Velasco et al. (2025), Inteligencia artificial y delincuencia organizada.
228 Europol. (2026, abril). Primer aniversario de la Operación Taskforce GRIMM: objetivos de violencia-como-servicio publicados en la lista de los más buscados de la UE. Agencia de la Unión Europea para la Cooperación Policial.
229 Entrevista con un investigador, (mayo de 2026)
230 Entrevista con un investigador, (mayo de 2026)
231 Entrevista con un investigador, (mayo de 2026)
232 Entrevista con un investigador, (mayo de 2026)
233 Entrevista con un investigador, (mayo de 2026)
234 Europol. (diciembre de 2025c).
235 Europol. (2025d).
236 Al Jazeera. (15 de diciembre de 2025). «¿Listos para matar?». Cómo las redes criminales en Suecia están reclutando a niños para matar.
237 Entrevista con un investigador, (mayo de 2026)
238 The Parliament Magazine. (4 de noviembre de 2025). Dentro de la guerra de bandas en Suecia: cómo las plataformas digitales están alimentando la violencia juvenil.
239 Europol. (1 de julio de 2025f).



RECOMENDACIONES

La IA ha reducido considerablemente el coste, la escala y el umbral de cualificación del reclutamiento criminal. Los marcos legales e institucionales existentes no fueron diseñados para esto. Las recomendaciones que siguen abordan tanto las deficiencias operativas inmediatas identificadas en este estudio como las reformas estructurales necesarias para comenzar a responder al reclutamiento asistido por IA. Ninguna intervención por sí sola podrá contener esta amenaza.

RECOMENDACIONES A CORTO PLAZO

RECOMENDACIÓN 1. DESIGNAR PUNTOS FOCALES PARA LA DELINCUENCIA ASISTIDA POR IA DENTRO DE LAS ESTRUCTURAS POLICIALES EXISTENTES

Los Estados miembros y los países socios deben designar puntos focales específicos dentro de las unidades existentes con responsabilidad específica en el reclutamiento criminal asistido por IA. Su función es garantizar que los casos con componentes de reclutamiento asistido por IA se señalen y compartan, en lugar de perderse en las bases de datos nacionales. Estos puntos focales deben estar conectados a los canales de comunicación bilaterales y multilaterales existentes de las fuerzas del orden.

RECOMENDACIÓN 2. ELABORAR UNA GUÍA DE DETECCIÓN PARA LOS PROFESIONALES

Debería elaborarse una guía operativa de detección para los profesionales de primera línea y difundirse a través de los canales de formación existentes para las fuerzas del orden y el apoyo a las víctimas. La guía debería:

- Describir cómo se presenta el reclutamiento asistido por IA en las principales tipologías de este estudio, utilizando ejemplos breves de casos y capturas de pantalla cuando sea posible.
- Destacar los indicadores a nivel de plataforma que merecen atención (ofertas de «trabajo en el extranjero» generadas algorítmicamente, traspasos repetidos a los mismos contactos de WhatsApp o Telegram, contacto a través de plataformas de juegos o de nicho, perfiles sintéticos de reclutadores).
- Proporcionar preguntas concretas para las entrevistas que permitan detectar la participación de la IA en los relatos de las víctimas, incluyendo preguntas sobre mensajes automatizados, herramientas de traducción, conversaciones con guion y paneles de control o interfaces de envío masivo de mensajes que las víctimas se vieron obligadas a utilizar.
- Explica por qué el contenido de reclutamiento generado por IA es cada vez más difícil de distinguir del material legítimo, y cómo reconocer la diferencia.
- Haga hincapié en que la detección depende del reconocimiento de patrones a nivel de cada caso (quién contactó con quién, con qué frecuencia, en qué plataformas y con qué traspasos), y no solo del análisis del contenido mensaje por mensaje.

RECOMENDACIÓN 3. ACTUALIZAR LOS PROTOCOLOS DE IDENTIFICACIÓN DE VÍCTIMAS PARA INCLUIR INDICADORES DE RECLUTAMIENTO ASISTIDO POR IA

Los equipos de primera línea no disponen actualmente de ningún protocolo para identificar el reclutamiento asistido por IA en los relatos de las víctimas. Los marcos de identificación de víctimas deben actualizarse para incluir indicadores de reclutamiento asistido por IA, tales como:

- Contacto en línea sostenido y de gran volumen que se mantiene a través de zonas horarias o plataformas, con respuestas rápidas, disponibilidad las 24 horas del día o comunicación simultánea desde múltiples «gestores».
- Ofertas de empleo u oportunidades de «trabajo en el extranjero» respaldadas por documentación digital y sitios web muy profesionales que parecen legítimos pero que no pueden verificarse de forma independiente, incluso cuando la documentación o los anuncios muestran signos de haber sido generados por IA.

- Relaciones románticas o basadas en la confianza que se desarrollan íntegramente en línea, a menudo a través de las fronteras, y que nunca dan lugar a un contacto en persona a pesar de una comunicación sostenida y emocionalmente intensa y de exigencias financieras o sexuales cada vez mayores.

- Reclutamiento iniciado a través de entornos de juego, plataformas especializadas o canales que no suelen asociarse con las rutas de la trata, incluyendo acercamientos por chat de voz y supuestos compañeros que utilizan avatares o voces sintéticas.

- Víctimas que denuncian que se les exigía manejar herramientas de IA (chatbots, paquetes de traducción, paneles de control de mensajes masivos) o gestionar un gran número de perfiles en línea como parte de su «trabajo», especialmente en entornos de centros de estafa o fraude.

RECOMENDACIÓN 4. ACLARAR LA APLICACIÓN DE LOS PRINCIPIOS DE NO PENALIZACIÓN A LA EXPLOTACIÓN FACILITADA POR LA IA

Las normas internacionales de no penalización para las personas víctimas de trata y explotación ya exigen que no se penalice a las víctimas por actos ilícitos cometidos como consecuencia directa de su explotación; debe aplicarse la misma lógica cuando dichos actos impliquen el manejo de herramientas de reclutamiento o fraude basadas en IA. Las víctimas se ven cada vez más obligadas a manejar herramientas de reclutamiento, fraude y comunicación basadas en IA dentro de centros de estafa y entornos relacionados con la delincuencia organizada. El conocimiento detallado de la infraestructura de reclutamiento basada en IA, los sistemas de envío masivo de mensajes, las identidades sintéticas o los paneles de traducción no debe considerarse prueba de participación voluntaria. Los marcos nacionales de identificación de víctimas deben aclarar explícitamente que, cuando el manejo de herramientas de IA forme parte de la situación de la víctima, se aplican los principios de no penalización existentes. El personal de primera línea, los fiscales y los investigadores deben recibir formación para reconocer este patrón y evitar clasificar erróneamente a las víctimas como delincuentes. Estas aclaraciones pueden integrarse en los mecanismos nacionales de derivación existentes en los Estados miembros de la UE y en los procedimientos equivalentes de las jurisdicciones de América Latina y el Caribe sin necesidad de nueva legislación.

RECOMENDACIÓN 5. ACTIVAR VÍAS DE INVESTIGACIÓN FINANCIERA SOBRE LAS REDES DE CAPTACIÓN ASISTIDAS POR IA

Las redes de captación asistidas por IA dependen de una infraestructura financiera rastreable: pagos en criptomonedas por herramientas de IA como servicio, acceso por suscripción a plataformas LLM ocultas y ganancias procedentes de operaciones de trata y fraude. Los flujos financieros suelen ser más recuperables que la infraestructura técnica. Un modelo de peso abierto alojado localmente no deja ningún registro del proveedor, pero las transacciones que lo financian pueden ser rastreables. Los puntos focales designados en virtud de la Recomendación 1 deben estar conectados con las unidades de inteligencia financiera y las redes de recuperación de activos existentes, con la orientación explícita de que los casos de reclutamiento asistido por IA deben evaluarse desde el principio en cuanto a su potencial para la investigación financiera. La base jurídica existe en los marcos AML existentes y en el artículo 29 de la Ley Modelo EL PACCTO.

RECOMENDACIONES A LARGO PLAZO

RECOMENDACIÓN 6. ESTABLECER UNA RECOPILACIÓN SISTEMÁTICA DE DATOS SOBRE EL RECLUTAMIENTO ASISTIDO POR IA

Actualmente, ninguna jurisdicción registra la participación de la IA en los casos de reclutamiento de forma sistemática o interoperable. Debería acordarse un conjunto de indicadores estandarizados que abarque el tipo de herramienta, la función dentro de la cadena de reclutamiento, la plataforma, el perfil de la víctima y la tipología delictiva, y que se incorpore a los requisitos de registro de casos en todas las jurisdicciones participantes. Paralelamente, debería desarrollarse un mecanismo de notificación estructurado para las organizaciones de la sociedad civil de primera línea, con el fin de recabar información sobre los casos de las organizaciones que se enfrentan al reclutamiento asistido por IA de forma operativa y con frecuencia antes que las fuerzas del orden.

RECOMENDACIÓN 7. REFORMAR LOS MARCOS LEGISLATIVOS PARA ABORDAR LAS LAGUNAS ESPECÍFICAS DEL RECLUTAMIENTO ASISTIDO POR IA

La legislación penal vigente no fue concebida para el reclutamiento asistido por IA y presenta lagunas que requieren atención legislativa:

- **Contenido de doble uso:** el material de reclutamiento generado por IA no es intrínsecamente ilegal. La reforma legislativa debería establecer el uso de la IA en una cadena de reclutamiento como factor agravante y definir las condiciones en las que constituye prueba de premeditación.

- **Modelos autohospedados:** los modelos de peso abierto, desplegados localmente y sin censura, no dejan registros de proveedores ni de acceso. Los marcos existentes crean un vacío de atribución que no se ha abordado.

La Ley Modelo EL PACCTO 2.0 sobre Inteligencia Artificial y Delincuencia ofrece disposiciones modelo para las jurisdicciones de América Latina y el Caribe en las tres áreas. Los Estados miembros de la UE deberían aprovechar el ciclo de implementación de la Ley de IA para abordar cualquier aspecto de la justicia penal que la Ley no cubra.

RECOMENDACIÓN 8. DESARROLLAR CAPACIDADES DE LA JUDICATURA Y LA FISCALÍA EN MATERIA DE DELITOS ASISTIDOS POR IA

Ya se está pidiendo a los tribunales que se pronuncien sobre las pruebas generadas por IA, la trazabilidad algorítmica y la intención delictiva en las cadenas de reclutamiento automatizadas sin disponer de la base técnica necesaria para hacerlo de forma fiable.[240,241] Deberían desarrollarse programas de formación que incluyan, como mínimo:

- Conocimientos técnicos básicos sobre el funcionamiento de los modelos generativos, los deepfakes y las herramientas de segmentación automatizada, y sobre los puntos débiles de las herramientas actuales de detección y atribución;

- Normas probatorias para las pruebas digitales generadas por IA;

- El marco de responsabilidad previsto en la Ley Modelo EL PACCTO y los instrumentos nacionales equivalentes;

240 Heaton, C., Cleary, S. y Navin, M. (24 de febrero de 2026). Las pruebas generadas por IA son una amenaza para la confianza pública en los tribunales. Centro Nacional de Tribunales Estatales.
241 Runyon, N. (8 de mayo de 2025). Los deepfakes a juicio: cómo los jueces abordan la autenticación de las pruebas generadas por IA. Instituto Thomson Reuters.

- Los problemas jurisdiccionales que plantean las redes de captación que abarcan varios países.

La Red Europea de Formación Judicial es el vehículo adecuado para los Estados miembros de la UE. Los canales de cooperación judicial existentes de EL PACCTO 2.0 y el mandato de desarrollo de capacidades recogido en el artículo 17, apartado 3, de la Ley Modelo proporcionan el mecanismo para las jurisdicciones de América Latina y el Caribe. Esto debe institucionalizarse, y no aplicarse de forma puntual, dado el ritmo al que evolucionan las capacidades de la IA y sus aplicaciones delictivas.

RECOMENDACIÓN 9. ESTABLECER OBLIGACIONES DE RENDICIÓN DE CUENTAS PARA LAS PLATAFORMAS Y REQUISITOS PARA LOS PROVEEDORES DE MODELOS DE IA EN RELACIÓN CON EL RECLUTAMIENTO DELICTIVO ASISTIDO POR IA

El reclutamiento delictivo analizado en esta investigación se lleva a cabo en plataformas generalistas y mediante IA comercial, que ahora proporcionan la mayor parte de la capacidad necesaria para industrializar ofertas de empleo engañosas, guiones de captación y reclutadores sintéticos. La IA reduce los umbrales de coste y de habilidades, permite un alcance multilingüe y permite a los grupos de delincuencia organizada segmentar a los usuarios vulnerables a gran escala, mientras que las salvaguardias existentes se centran en contenidos individuales en lugar de en los embudos de reclutamiento.

- **Plataformas.** Las redes sociales, las plataformas de mensajería y los juegos proporcionan la infraestructura principal para el reclutamiento asistido por IA; sus sistemas de recomendaciones, publicidad y mensajería se explotan habitualmente para llegar a usuarios vulnerables y captarlos, al tiempo que se elude la moderación tradicional de contenidos. Por lo tanto, las obligaciones reglamentarias deben centrarse en los embudos de reclutamiento y los patrones de explotación, y no solo en publicaciones nocivas aisladas. Las medidas recomendadas incluyen:

  - Publicar informes de transparencia sobre la aplicación de la ley en materia de reclutamiento (ofertas de empleo falsas, «trabajo en el extranjero», sextorsión, captación para cárteles y bandas, OCSEA), desglosados por tipología, grupo de edad y corredor geográfico.

  - Realizar y documentar evaluaciones de riesgo sistémico sobre cómo las funciones de recomendación, búsqueda y descubrimiento de contactos amplifican los contenidos de reclutamiento engañosos, prestando especial atención a los menores y a los usuarios económicamente vulnerables.

  - Detectar e interrumpir los canales de reclutamiento mediante la agregación de señales en contenidos y funciones (plantillas de empleo reutilizadas, traspasos repetidos fuera de la plataforma, números de teléfono compartidos, vías de pago, rutas de viaje e invitaciones a grupos).

  - Participar en el intercambio de inteligencia entre plataformas centrado en los indicadores de explotación por parte del crimen organizado (marcas de centros de estafa, alias de agencias de reclutamiento, guiones comunes, carteras de criptomonedas y datos de contacto), con especial atención a la explotación humana.

  - Incorporar requisitos de protección infantil y prevención de la explotación en el diseño de productos para feeds, mensajes directos, funciones de grupo y chat dentro de los juegos, incluyendo ajustes predeterminados adecuados a la edad y detección de riesgos de captación.

  - Garantizar canales de escalamiento accesibles y dotados de recursos suficientes para que los denunciantes de confianza (líneas directas, unidades especializadas de las fuerzas del orden, ONG) puedan informar de presuntas campañas de reclutamiento y obtener respuestas oportunas y significativas.

- **Proveedores de modelos de IA:** Los grandes modelos de lenguaje comerciales ya generan anuncios de empleo, contratos, guiones románticos y perfiles sintéticos realistas que pueden ser reutilizados para el reclutamiento por parte de organizaciones criminales de formas que, a menudo, son indistinguibles, si se analizan de forma aislada, del contenido legítimo. Por lo tanto, la detección debe centrarse en el comportamiento y los patrones, no en mensajes o indicaciones aisladas. Las medidas recomendadas incluyen:

  - Desarrollar la detección de comportamientos específicos del reclutamiento (patrones de empleos falsos, ofertas de «trabajo en el extranjero», embudos de captación y sextorsión, puestos de «atención al cliente» al estilo de los centros de estafa). Estos deben basarse en indicadores actualizados para captar los embudos de reclutamiento modernos impulsados por la IA.

  - Agregar señales de riesgo entre sesiones y cuentas (plantillas de empleo reutilizadas, puntos de contacto compartidos, traspasos repetidos fuera de la plataforma, rutas comunes de viaje o de pago) en lugar de basarse en el filtrado de mensajes aislados.

  - Desarrollar modelos de «equipo rojo» específicamente para la trata de personas, el trabajo forzoso, los centros de estafa, el blanqueo de dinero, la violencia como servicio y los escenarios de OCSEA, e implementar medidas de mitigación concretas.

  - Aplicar procedimientos graduados de «conozca a su cliente» y de verificación para los niveles de API de mayor riesgo (generación masiva, ajuste fino, herramientas autónomas o de agente), al tiempo que se mantiene un bajo nivel de fricción para el uso ordinario y de bajo riesgo.

  - Proporcione canales estructurados de telemetría y denuncia de abusos para que organizaciones de confianza (líneas directas, unidades especializadas de las fuerzas del orden, ONG) puedan señalar posibles canales de reclutamiento utilizando sus modelos.

  - Limitar o aislar en un entorno de pruebas las funciones que reducen sustancialmente el coste de las operaciones de reclutamiento industrial (envío masivo de mensajes, generación automatizada de clientes potenciales, creación de perfiles en gran volumen y sembrado de cuentas).

  - Incorporar de forma predeterminada requisitos de protección infantil y prevención de la explotación en todos los productos de «asistencia laboral», de búsqueda de empleo y de herramientas para desarrolladores, incluyendo un diseño adecuado a la edad y controles de riesgo de captación.

En conjunto, estas obligaciones exigen que las plataformas saquen a la luz los canales de reclutamiento del crimen organizado y que los proveedores de IA detecten y restrinjan los comportamientos impulsados por modelos que los alimentan, sentando las bases de una defensa coordinada en toda la cadena de explotación. Esto no detendrá la amenaza en constante evolución, pero comenzará a exponer y restringir el uso de herramientas comerciales abiertas que actualmente facilitan la explotación humana.

CONCLUSIONES

Las pruebas examinadas en este estudio respaldan una conclusión clara: la IA ha cambiado lo que es operativamente posible en el reclutamiento criminal para la explotación humana. La lógica central no ha cambiado. Las redes criminales necesitan llegar a posibles reclutas, seleccionar objetivos viables, establecer credibilidad y evitar la detección. La IA reduce el coste y la fricción de cada una de esas funciones simultáneamente, al tiempo que aumenta su precisión y escala. El resultado es un proceso de reclutamiento cada vez más profesional e industrializado.

Las herramientas de IA reducen las barreras de entrada para los nuevos actores delictivos, al tiempo que permiten a las redes ya establecidas ampliar su capacidad sin que ello suponga un aumento proporcional de los costes, la plantilla o el riesgo operativo. Un solo operador que utilice herramientas y plataformas comerciales de IA puede alcanzar un alcance que antes habría requerido una plantilla humana multilingüe y con una logística compleja. Este cambio estructural es significativo independientemente de las herramientas específicas que se utilicen.

Hay tres características estructurales del panorama actual que merecen especial atención. En primer lugar, la mayor parte del reclutamiento delictivo asistido por IA se lleva a cabo en infraestructuras comerciales. Los algoritmos de recomendación de las plataformas, los modelos de lenguaje grandes (LLM) convencionales, las suites de traducción con IA y las herramientas de medios sintéticos no son productos delictivos: son servicios comerciales legítimos que los actores delictivos explotan. Esto significa que las contramedidas centradas exclusivamente en las herramientas de IA delictivas solo abordarán los casos marginales. El principal desafío es que gran parte del contenido de reclutamiento delictivo producido con herramientas comerciales es, de por sí, técnicamente indistinguible del contenido legítimo, y no activa los sistemas de moderación de contenido diseñados para detectarlo.

EL PACCTO 2.0

En segundo lugar, las redes delictivas no se dirigen a los niños y adolescentes a pesar de su vulnerabilidad; se dirigen a ellos precisamente por ella. La menor responsabilidad legal, la accesibilidad a las plataformas, la manipulabilidad psicológica y los puntos de entrada basados en los pares a través de las redes sociales son ventajas operativas que el reclutamiento delictivo explota explícitamente. Los deepfakes de IA, la clonación de voz y los sistemas de captación basados en el aprendizaje por refuerzo están calibrados específicamente en función de cómo se comunican los objetivos más jóvenes, en qué confían y dónde son más débiles sus defensas.

En tercer lugar, la IA ha hecho que el reclutamiento sea más eficaz para un grupo diverso de actores delictivos. Las organizaciones documentadas en este estudio abarcan desde grupos de explotación en línea vagamente interconectados hasta empresas multinacionales de centros de estafa que gestionan a cientos de trabajadores a través de las fronteras; desde redes de tráfico de drogas a pie de calle hasta cárteles transnacionales con una logística de nivel militar. Se diferencian en su escala, estructura, recursos, capacidad coercitiva y propósito delictivo. Su enfoque de la IA y el uso de herramientas específicas de IA también es exclusivo de cada una de ellas y puede variar entre organizaciones o incluso entre actores delictivos individuales. Sin embargo, lo que comparten es que la IA ha hecho que el reclutamiento sea más eficaz para cada una de ellas.

La trayectoria de la amenaza que supone el reclutamiento impulsado por la IA tiene dos dimensiones. La primera es la automatización. El reclutamiento totalmente impulsado por la IA, en el que los sistemas automatizados llevan a cabo la selección de objetivos, la captación y el fomento de la confianza sin supervisión humana, no es un escenario futuro. Se trata de un modelo operativo que está en sus inicios, pero que ya es creíble. A medida que los sistemas de IA con capacidad de agencia se vuelvan más accesibles y fiables, las restricciones humanas restantes disminuirán aún más. La segunda dimensión es menos visible, pero igualmente significativa: la IA cambia no solo la eficiencia del reclutamiento criminal, sino potencialmente su carácter. El desarrollo de estrategias impulsadas por la IA puede generar, probar y perfeccionar enfoques de reclutamiento en contextos lingüísticos, culturales y demográficos con una profundidad que supera la capacidad humana por sí sola. Esto significa que las operaciones de reclutamiento criminal pueden desarrollar tácticas que difieran de los patrones establecidos, produciendo nuevas formas de explotación, dirigiéndose a poblaciones que antes no estaban en riesgo y generando enfoques que las fuerzas del orden aún no tienen marcos para reconocer. Es posible que la amenaza que surge no se parezca a la amenaza que se está vigilando. Ambas dimensiones reducen el margen de maniobra para elaborar respuestas eficaces.

Las recomendaciones de este informe reflejan esta evaluación. Las personas al final de este proceso no son puntos de datos. Son niños preparados para la explotación, adultos engañados para realizar trabajos forzados e individuos obligados a desempeñar roles delictivos. La IA no ha cambiado lo que se les hace. Ha hecho que el engaño sea más difícil de reconocer, el reclutamiento más convincente y el daño más difícil de prevenir.

BIBLIOGRAFÍA

Aguilar Antonio, J.M. (2025). «Uso de la inteligencia artificial por parte de redes criminales de alto riesgo». Expertise France. DOI: 10.5281/zenodo.16750778

Aldawsari, A. H. (2024), «Evaluación de herramientas de traducción: Google Translate, Bing Translator y Bing AI en coloquialismos árabes», Arab World English Journal (AWEJ), número especial sobre ChatGPT, pp. 237-251.

Al Jazeera. (15 de diciembre de 2025). «¿Listo para matar?». Cómo las redes criminales en Suecia están reclutando a niños para matar.

Amerhauser, K. y Goodwin, A. (2026), «Un mundo de engaños: Mapeo del panorama del fenómeno de los centros de estafa globales», Iniciativa Global contra el Crimen Organizado Transnacional.

Andrews, C. (2025), «La Comisión Europea retira de su agenda la Directiva sobre responsabilidad en materia de IA», Asociación Internacional de Profesionales de la Privacidad.

Ang, A. (2025), «La IA podría sustituir a las personas en los complejos de estafas del sudeste asiático, lo que podría socavar los esfuerzos para detenerlos», Fortune, 20 de noviembre.

Aoukar, Y., Péron, C., Maddox, L. y Apple, L. (2024), «IA generativa: una nueva amenaza para la explotación y el abuso sexual infantil en línea», Fundación Bracket, Value for Good y Centro de Inteligencia Artificial y Robótica del UNICRI.

Associated Press (2025), «El jefe de seguridad de México afirma que los cárteles de la droga están reclutando a exsoldados colombianos», Military.com, 11 de junio.

Bender, E. M., Gebru, T., McMillan-Major, A. y Shmitchell, S. (2021), «On the dangers of stochastic parrots: Can language models be too big?», Actas de la Conferencia ACM 2021 sobre Equidad, Responsabilidad y Transparencia, pp. 610-623.

Bledsoe, R. (2025), «El papel de las redes sociales en el reclutamiento de cárteles», Centro de Estudios Estratégicos e Internacionales.

Burgess, M. (16 de marzo de 2026). Las modelos se presentan para ser el rostro de las estafas de IA. WIRED.

Centro Canadiense para la Protección de la Infancia (2024), «La policía y la agencia de protección infantil afirman que los padres deben estar al tanto de los deepfakes de IA con contenido sexual explícito», Centro Canadiense para la Protección de la Infancia.

Conrad, C., Heintz, E. y Mace, M. (2025). Un posible nexo entre las estafas forzadas y la sextorsión de menores con fines económicos. International Justice Mission.

EL PACCTO 2.0

Chaudhari, S., Aggarwal, P., Murahari, V., Rajpurohit, T., Kalyan, A., Narasimhan, K., Deshpande, A. y Castro da Silva, B. (2025), «RLHF deciphered: A critical analysis of reinforcement learning from human feedback for LLMs», ACM Computing Surveys, 58(2).

Chu, M. D., Gerard, P., Pawar, K., Bickham, C. y Lerman, K. (2025), «Illusions of intimacy: Emotional attachment and emerging psychological risks in human-AI relationships», arXiv.

Collier, K. (2026), «La crisis de la explotación infantil por IA ya está aquí», NBC News.

Colman, B. (2026), «Los deepfakes en tiempo real están reescribiendo las reglas de la seguridad infantil», Foro Económico Mundial, 30 de abril.

Comolli, V. (2025), «Explotación sexual infantil organizada: Abordar los motivos y las necesidades de respuesta en el Sudeste Asiático», Iniciativa Global contra la Delincuencia Organizada Transnacional.

Cox, J. (2026), «"HELLO BOSS": Dentro del software chino de deepfakes en tiempo real que impulsa estafas en todo el mundo», 404 Media.

CyberScoop (2024a), «Las autoridades federales investigan a 764, un grupo en línea que se dedica a cometer delitos cibernéticos y violentos contra menores**», CyberScoop.**

Dalby, C. (2021), «Cómo los cárteles mexicanos utilizan los videojuegos para reclutar a niños**», InSight** Crime.

Di Girolamo, M. (2026), «Engañoso por diseño: Las herramientas basadas en IA que alimentan la industria de las estafas», C4ADS.

Duffy, C. (2024), «"No hay barreras de seguridad". Esta madre cree que un chatbot de IA es responsable del suicidio de su hijo», CNN, 30 de octubre.

Comisión Europea (2025), **«**Comunicación sobre un Plan de Acción de la UE contra el tráfico de drogas y la delincuencia organizada», COM (2025) 744 final.

Europol (2024a), «La operación europea contra los muleros de dinero da lugar a 1 803 detenciones», Europol.

Europol (2024b), «Evaluación de la amenaza de la delincuencia organizada en Internet (IOCTA) 2024», Agencia de la Unión Europea para la Cooperación Policial.

Europol (2024c), «El rastro documental acaba en penas de cárcel para 1013 mulas de dinero», Europol.

Europol (2025a), «El ADN cambiante de la delincuencia grave y organizada: Evaluación de la amenaza de la delincuencia grave y organizada en la UE 2025 (EU-SOCTA 2025)», Agencia de la Unión Europea para la Cooperación Policial.

Europol (2025b), «Operación global contra Kidflix, una importante plataforma de explotación sexual infantil con casi dos millones de usuarios», Europol.

Europol (2025c), «Grupo de trabajo operativo GRIMM: 193 detenciones en 6 meses contra las redes de violencia como servicio», Europol.

Europol (2025d), «Grupo de trabajo operativo GRIMM: Lucha contra la violencia como servicio y el reclutamiento de jóvenes delincuentes para la delincuencia grave y organizada», Europol.

Europol (2025e), «Robar, vender y repetir: cómo los ciberdelincuentes comercian y explotan tus datos — Evaluación de la amenaza de la delincuencia organizada en Internet (IOCTA) 2025», Oficina de Publicaciones de la Unión Europea.

Europol (2025f), «Adolescentes reclutados como sicarios: Dinamarca y Suecia contraatacan la violencia como servicio», Europol.

Europol (2026), «IOCTA 2026 — El panorama cambiante de las amenazas: cómo el cifrado, los proxies y la IA están ampliando la ciberdelincuencia», Europol.

Europol. (abril de 2026). «Primer aniversario de la operación Taskforce GRIMM: objetivos de violencia-como-servicio publicados en la lista de los más buscados de la UE.». Agencia de la Unión Europea para la Cooperación Policial.

Europol, UNICRI y Trend Micro (2020), «Usos maliciosos y abusos de la inteligencia artificial», Europol.

Eurostat (2026a), «9 678 víctimas de la trata de personas registradas en 2024», Comisión Europea.

Eurostat (2026b), «Estadísticas sobre la trata de seres humanos», Comisión Europea.

Oficina Federal de Investigación (2023), «Actores maliciosos que manipulan fotos y vídeos para crear contenido explícito y esquemas de sextorsión», Centro de Denuncias de Delitos en Internet.

Oficina Federal de Investigación (2025a), «Las redes violentas en línea se dirigen a poblaciones vulnerables y menores de edad en todo Estados Unidos y en todo el mundo», Centro de Denuncias de Delitos en Internet, Alerta PSA I-030625-PSA.

Oficina Federal de Investigación (2025b), «The Com: El robo, la extorsión y la violencia son una amenaza creciente para los jóvenes en Internet», Centro de Denuncias de Delitos en Internet, Alerta n.º I-072325-3-PSA.

Felbab-Brown, V. (2026), «Cómo utilizan los cárteles mexicanos los drones, ahora y en el futuro», Brookings Institution, 18 de febrero.

FinCEN (2025), «FinCEN publica un aviso sobre la sextorsión con fines económicos», Departamento del Tesoro de EE. UU., Red de Control de Delitos Financieros, 8 de septiembre.

Fox News (2025), «El FBI investiga a más de 350 personas vinculadas a la violenta red online "764"», Fox News, 12 de diciembre.

EL PACCTO 2.0

Iniciativa Global contra el Crimen Organizado Transnacional (2025), «Niños soldados en Europa: ¿Por qué el crimen organizado recluta cada vez más a menores?», Iniciativa Global contra el Crimen Organizado Transnacional.

González, F. (2026), «Cómo el cártel de la droga mexicano "CJNG" adoptó la IA, los drones y las redes sociales», WIRED.

Gressel, G., Pankajakshan, R., Rozenfeld, S., Li, L., Franceschini, I., Achuthan, K. y Mirsky, Y. (2025), «Amor, mentiras y modelos de lenguaje: investigación sobre el papel de la IA en las estafas sentimentales», preimpresión de arXiv arXiv:2512.16280.

Heaton, C., Cleary, S. y Navin, M. (2026), «Las pruebas generadas por IA son una amenaza para la confianza pública en los tribunales», Centro Nacional de Tribunales Estatales.

Henderson Vaughan, E. (2026), «El trabajo nunca se detiene: un primer vistazo a los datos de 2025 del NCMEC», Centro Nacional para Niños Desaparecidos y Explotados, 31 de marzo.

Horwitz, J. (2025), «El coqueto chatbot con IA de Meta invitó a un jubilado a Nueva York. Nunca regresó a casa», Reuters, 14 de agosto.

InSight Crime (2025), «¿Está América Latina preparada para mitigar el reclutamiento digital por parte de los grupos delictivos?», InSight Crime. Disponible en:

Instituto para el Diálogo Estratégico (2025), «Redes de daño: un recurso informativo centrado en las víctimas sobre la red de sextorsión 764», Instituto para el Diálogo Estratégico.

Fundación Internet Watch (2025a), «Chatbots con IA y abuso sexual infantil: una llamada de atención para adoptar medidas de protección urgentes», Fundación Internet Watch.

Internet Watch Foundation (2025b), «Imágenes generadas por IA», en Informe anual de datos y análisis 2025, Internet Watch Foundation.

Internet Watch Foundation (2025c), «Extorsión con coacción sexual», en Informe anual de datos y análisis 2025, Internet Watch Foundation.

Internet Watch Foundation (2025d), «Tomar medidas para ayudar a los adolescentes mayores», en Informe anual de datos y análisis 2025, Internet Watch Foundation.

Fundación Internet Watch (2026), «Daño sin límites: material de abuso sexual infantil generado por IA a través de los ojos de nuestros analistas», Fundación Internet Watch.

INTERPOL (2023), «América: detenidos 257 presuntos traficantes de migrantes y de personas», INTERPOL.

INTERPOL (2024a), «Dentro de la investigación de INTERPOL sobre la trata de personas facilitada por Internet», INTERPOL.

INTERPOL (2024b), «Evaluación de INTERPOL sobre el fraude financiero: una amenaza mundial impulsada por la tecnología», INTERPOL.

Kapko, M. (2025), «Detenidos y acusados los líderes de 764, una red mundial de sextorsión infantil», CyberScoop, 1 de mayo.

KELA Cyber Threat Intelligence (2025), «Informe sobre amenazas de IA de 2025: Cómo los ciberdelincuentes están utilizando la tecnología de IA como arma», KELA Cyber Threat Intelligence.

Kovacs, E. (2025), «WormGPT 4 y KawaiiGPT: los nuevos LLM oscuros impulsan la automatización de la ciberdelincuencia», SecurityWeek, 25 de noviembre.

Levesque, J. (2025), «La amenaza del tráfico de personas potenciada por la IA: análisis de la evolución tecnológica de las operaciones de tráfico de personas con herramientas de IA», Journal of Human Trafficking.

Lin, Z., Li, Z., Liao, X. y Wang, X. (2025), «Consiglieres in the shadow: Understanding the use of uncensored large language models in cybercrimes», preimpresión de arXiv arXiv:2508.12622.

Malwarebytes. (23 de marzo de 2026). «Las estafas recurren a «modelos de IA» para cerrar el trato en videollamadas deepfake». Malwarebytes Labs.

Mascellino, A. (2025), «La IA Xanthorox de la Darknet ofrece herramientas personalizables para hackers», Infosecurity Magazine.

McPherson, P. (15 de septiembre de 2025). «ChatGPT se utilizó «para ayudar a los estafadores a hacer de las suyas» en un complejo de fraude asiático». Reuters.

Montes Delijorge, J. (2026), «La IA está transformando el panorama de la explotación humana: un análisis de las amenazas en evolución», Medium.

Centro Nacional para Niños Desaparecidos y Explotados (2025), «El NCMEC publica nuevos datos: 2024 en cifras», Centro Nacional para Niños Desaparecidos y Explotados.

Centro Nacional de Innovación, Tecnología y Educación Antiterrorista (2025), «El enjuiciamiento del 764: un análisis de los cargos federales y estatales», Universidad de Nebraska Omaha.

NETSCOUT (2025), «DDoS por encargo y la evolución del uso de la IA», NETSCOUT.

News 4 Tucson KVOA-TV (s. f.), «Una madre de Arizona con un mensaje para los padres después de que su hija fuera víctima de un catfishing», KVOA.

Noyes, D. (2025), «Amenaza "764": el FBI afirma que hay depredadores que se fijan en niños a través de videojuegos y redes sociales en el Área de la Bahía y más allá», ABC7 News, 6 de mayo.

Oficina del Representante Especial y Coordinador de la OSCE para la Lucha contra la Trata de Seres Humanos (2026), «La trata de personas para operaciones de estafa cibernética: Implicaciones para la región de la OSCE — Evaluación exploratoria», Organización para la Seguridad y la Cooperación en Europa.

Oficina del Alto Comisionado de las Naciones Unidas para los Derechos Humanos (2026), «Un «problema complejo»: en busca de soluciones basadas en los derechos humanos para la trata de personas destinada a operaciones de estafa cibernética en el sudeste asiático», Naciones Unidas.

OpenAI (2023), «Informe técnico sobre GPT-4», arXiv:2303.08774.

OpenAI. (febrero de 2026). «Disrupting malicious uses of our models: An update».

Organización para la Seguridad y la Cooperación en Europa y Oficina de Apoyo Regional del Proceso de Bali (2024), «Nuevas fronteras: El uso de la inteligencia artificial generativa para facilitar la trata de personas», OSCE.

ParentsTogether Action y HEAT Initiative (2025), «Cariño, por favor, vuelve pronto: Explotación sexual, manipulación y violencia en las cuentas infantiles de Character AI», ParentsTogether Action.

Pearl, M., Brock, J. y Kumar, A. (2025), «Delving into the dangers of DeepSeek», Centro de Estudios Estratégicos e Internacionales, 24 de febrero.

Railway Children (2023), «El rostro cambiante de las redes de tráfico de drogas interregionales: por qué es importante y qué estamos haciendo al respecto», Railway Children.

Remote Work Europe (s. f.), «Guía sobre estafas en el trabajo a distancia», Remote Work Europe.

Ríos Morales, H. (2025), «TikTok se ha convertido en la principal plataforma que utilizan los cárteles para reclutar adolescentes en México, según un informe», TechTimes, 4 de diciembre.

Real Policía Montada de Canadá (2024), «Grupos violentos en línea que explotan a niños y jóvenes», Real Policía Montada de Canadá.

Runyon, N. (2025), «Deepfakes a juicio: cómo los jueces abordan la autenticación de pruebas de IA», Instituto Thomson Reuters.

Schultz, J. (2025), «Abuso de los modelos de lenguaje a gran escala por parte de los ciberdelincuentes», Cisco Talos, 25 de junio.

Seminario sobre Violencia y Paz, El Colegio de México, Laboratorio de Odio y Concordia y Civic A.I. Lab (2025), «Reclutamiento criminal en TikTok: una investiga a documenta más de 100 cuentas activas en México».

Signa Lab ITESO (2024), «TikTok y el reclutamiento para actividades criminales en México: Hallazgos preliminares», El Colegio de México, abril.

Simonovich, V. (2025), «Investigación sobre amenazas de Cato CTRL: variantes de WormGPT impulsadas por Grok y Mixtral», Cato Networks, 17 de junio.

Srinath, C. y Srivastava, S. (2024), «Rastreo web dinámico impulsado por IA GEN en Azure utilizando Automation Account y GPT-3.5», International Journal of Engineering and Advanced Technology.

Swanson, E. (2025), «El modelo de IA DeepSeek-R1 es 11 veces más propenso a generar contenido perjudicial, según un estudio de seguridad», Enkrypt AI.

The Parliament Magazine. (4 de noviembre de 2025). «Dentro de la guerra de bandas en Suecia: cómo las plataformas digitales están alimentando la violencia juvenil».

Tynan, R. (2026), «Cómo los cárteles mexicanos están utilizando contenido generado por IA para modernizar el reclutamiento», Alice, Medium.

Departamento de Seguridad Nacional de EE. UU. (2024), «Impacto de la inteligencia artificial (IA) en las actividades delictivas e ilícitas», Programa de Intercambio Analítico Público-Privado.

Departamento de Justicia de EE. UU. (2025a), «Líderes de la red 764 detenidos y acusados de dirigir una red global de explotación infantil», Oficina de Asuntos Públicos.

Departamento de Justicia de EE. UU. (2025b), «Líder de Arizona de la red extremista violenta "764" acusado de dirigir una red de explotación infantil, apoyar a terroristas, producir y distribuir pornografía infantil y otros delitos», Oficina de Asuntos Públicos.

Departamento de Justicia de EE. UU. (2025c), «El líder del grupo extremista "764" se declara culpable de los cargos de violación de la ley RICO y de explotación infantil», Oficina de Asuntos Públicos.

Parlamento del Reino Unido (2025), «Tráfico de drogas en las redes de condados», Hansard, debate en la Cámara de los Lores, 24 de febrero.

UNICEF (2026), «"El abuso mediante deepfakes es abuso", advierte UNICEF», UNICEF Canadá.

Oficina de las Naciones Unidas contra la Droga y el Delito (2024), «Informe mundial sobre la trata de personas 2024», UNODC.

Oficina de las Naciones Unidas contra la Droga y el Delito (2025), «Punto de inflexión: Implicaciones globales de los centros de estafa, la banca clandestina y los mercados ilícitos en línea en el sudeste asiático», UNODC.

Velasco, C., García Periche, J., Gómez Gómez, J. D. y Bueno Benedí, M. (2025), «Inteligencia artificial y delincuencia organizada», EL PACCTO 2.0, Expertise France y FIAP.

Velasco, C., Rodríguez, F. A., Bueno, B. M. y Cassuto, T. (2025). «La Instrumentalización de la Inteligencia Artificial: cómo la IA reconfigura el mundo». EL PACCTO 2.0, Expertise France y FIIAPP.

EL PACCTO 2.0

EL PACCTO 2.0

EU-LAC Partnership on justice and security