pp. 94-133
Recibido: 17-04-2026
Aceptado: 29-05-2026
Publicado: 20-06-2026

Inteligencia Artificial y Conductas Ilícitas Emergentes en Entornos Digitales: Desafíos para el Derecho Penal y la Libertad de Expresión en Ecuador

Artificial Intelligence and Emerging Illicit Conduct in Digital Environments: Challenges for Criminal Law and Freedom of Expression in Ecuador

Henry Antonio Armijos-Campoverde[1]
Docente
Universidad Técnica Particular de Loja [UTPL] -Ecuador-
haarmijos@utpl.edu.ec

Resumen

El avance de la inteligencia artificial generativa ha transformado los entornos digitales y ampliado la producción automatizada de contenidos, generando nuevos desafíos para el derecho penal, la protección de datos personales y la libertad de expresión. En ese contexto la presente investigación analiza las principales conductas ilícitas emergentes asociadas al uso de inteligencia artificial en Ecuador; especialmente, aquellas vinculadas con suplantación de identidad digital, manipulación informativa automatizada y fraude informático, examinando su posible adecuación a los tipos penales vigentes y sus tensiones con los derechos fundamentales. La investigación se desarrolla mediante un enfoque cualitativo de carácter jurídico-

1 Docente No Titular de la Universidad Técnica Particular de Loja. Magister en Derecho penal. Mención en Derecho Procesal Penal. Magister en Derecho Civil y Procesal Civil por la Universidad Técnica Particular de Loja. Abogado por la Universidad Técnica Particular de Loja. Doctorado en ciencias sociales, criminológicas y del comportamiento. Integrante del Grupo de Investigación jurídica aplicada - GRIJAX. **ORCID**: https://orcid.org/0000-0003-0888-6013

doctrinal sustentado en el análisis de la Constitución de la República del Ecuador [CRE] (2008), el Código Orgánico Integral Penal [COIP] (2024), la Ley Orgánica de Protección de Datos Personales [LOPDP] (2022) y doctrina especializada. Los resultados evidencian que el ordenamiento jurídico ecuatoriano permite abordar parcialmente determinadas conductas derivadas del uso indebido de inteligencia artificial; sin embargo, persisten vacíos regulatorios y riesgos de sobrecriminalización. Se concluye que la respuesta estatal debe garantizar la protección de bienes jurídicos sin afectar desproporcionadamente la libertad de expresión ni el principio de legalidad penal.

Palabras clave: Inteligencia artificial, derecho penal, libertad de expresión, desinformación digital, protección de datos personales.

Abstract

The advanceof generative artificial intelligence has transformed digital environments and expanded the automated production of content, creating new challenges for criminal law, personal data protection, and freedom of expression. In this context the study analyzes the main emerging unlawful conduct associated with the use of artificial intelligence in Ecuador, particularly those related to digital identity theft, automated information manipulation, and cyber fraud, examining its possible classification under existing criminal offenses and the tensions with fundamental rights. The research is conducted using a qualitative legal-doctrinal approach based on the analysis of the Constitution of the Republic of Ecuador (2008), the Comprehensive Organic Criminal Code (2024), the Organic Law on Personal Data Protection (2022), andspecialized legal doctrine. The

Recibido: 17-04-2026
Aceptado: 29-05-2026
Publicado: 20-06-2026

results show that the Ecuadorian legal system allows for the partial regulation of certain forms of conduct arising from the misuse of artificial intelligence; however, regulatory gaps and risks of overcriminalizationpersist. It is concluded that the state responses must ensure the protection of legal interests without unduly restricting freedom of expression or the principle of legality in criminal law.

Keywords: artificial intelligence, criminal law, freedom of expression, digital disinformation, personal data protection.

1. Introducción

El desarrollo de la inteligencia artificial generativa ha transformado significativamente los procesos de producción, circulación y consumo de contenidos en entornos digitales. Las herramientas basadas en modelos algorítmicos permiten generar textos, imágenes, audios y videos sintéticos con altos niveles de realismo, ampliando las posibilidades de automatización comunicacional. Sin embargo, estas tecnologías también han facilitado la aparición de prácticas potencialmente lesivas relacionadas con suplantación de identidad, manipulación informativa, fraude informático y difusión masiva de contenidos alterados, generando nuevos desafíos para los sistemas jurídicos contemporáneos (Castells, 2009; Chesney & Citron, 2019).

La expansión de contenidos sintéticos y mecanismos automatizados de comunicación ha modificado las dinámicas tradicionales de circulación de información en plataformas digitales. De acuerdo con Van Dijck (2016), las plataformas operan mediante arquitecturas algorítmicas capaces de influir en la visibilidad, distribución y consumo de contenidos, incrementando el impacto potencial de

prácticas desinformativas o manipulativas. En este contexto, fenómenos como los *deepfakes*, la automatización de perfiles digitales y la utilización de inteligencia artificial con fines engañosos plantean dificultades relacionadas con la atribución de responsabilidad jurídica y la identificación de conductas susceptibles de afectar bienes jurídicos protegidos.

Desde la perspectiva jurídico penal, estos fenómenos generan interrogantes complejos respecto de la adecuación típica de determinadas conductas mediadas por inteligencia artificial. La utilización de sistemas automatizados para producir contenidos falsificados, alterar información o inducir engaños digitales obliga a analizar si el ordenamiento jurídico vigente permite responder eficazmente frente a nuevas formas de afectación patrimonial, vulneración de la intimidad, uso indebido de datos personales o manipulación de procesos comunicacionales.

No obstante, la expansión de mecanismos sancionatorios frente a la desinformación o la automatización comunicacional también plantea riesgos de sobrecriminalización incompatibles con el principio de mínima intervención penal y con las garantías derivadas de la libertad de expresión (Balkin, 2018).

En el caso ecuatoriano, el análisis adquiere especial relevancia debido a la coexistencia de normas destinadas a proteger bienes jurídicos vinculados con intimidad, patrimonio, datos personales y seguridad informática, junto con garantías constitucionales relacionadas con libertad de expresión, acceso a la información y legalidad penal.

La Constitución de la República del Ecuador [CRE] (2008) reconoce el derecho a buscar, recibir, intercambiar y difundir información por cualquier medio, así como la protección de datos personales y la tutela de la intimidad (arts. 66 y 92). Asimismo, el Código Orgánico Integral Penal [COIP] (2025) tipifica: conductas relacionadas con estafa (art. 186); revelación ilegal de bases de datos (art. 229); acceso no consentido a un sistema informático (art. 230); y, revelación de secretos o información personal de terceros. De igual manera, la Ley Orgánica de Protección de Datos Personales [LOPDP] (2021) establece principios orientados a garantizar el tratamiento legítimo de datos personales en entornos digitales.

En virtud de ello, cualquier respuesta estatal dirigida a enfrentar prácticas digitalmente lesivas derivadas del uso de inteligencia artificial debe observar criterios de necesidad, proporcionalidad y estricta legalidad. La Corte IDH (2010) ha sostenido que la libertad de expresión constituye uno de los pilares esenciales de las sociedades democráticas y que las restricciones estatales deben interpretarse restrictivamente para evitar afectaciones indebidas al debate público. Por ello, la regulación de fenómenos vinculados con desinformación automatizada, contenidos sintéticos o manipulación digital exige distinguir entre conductas que generan una afectación jurídicamente relevante y aquellas que permanecen protegidas por el ejercicio legítimo de derechos fundamentales.

La presente investigación tiene como objetivo analizar las principales conductas ilícitas emergentes asociadas al uso de inteligencia artificial en entornos

digitales y examinar en qué medida el ordenamiento jurídico ecuatoriano permite abordarlas mediante las figuras normativas actualmente vigentes. El estudio sostiene que, aunque determinadas conductas mediadas por inteligencia artificial pueden subsumirse parcialmente en tipos penales existentes, persisten dificultades de adecuación típica respecto de prácticas como la creación y difusión de *deepfakes* sin finalidad patrimonial, la manipulación informativa automatizada y determinadas formas de desinformación digital, las cuales plantean desafíos regulatorios relevantes. Asimismo, se identifican riesgos de expansión punitiva que exigen una interpretación compatible con los derechos fundamentales y con el principio de legalidad penal.

2. Método

La presente investigación se desarrolló mediante un enfoque cualitativo de carácter jurídico-doctrinal orientado al análisis de las conductas ilícitas emergentes derivadas del uso de inteligencia artificial en entornos digitales y su incidencia en el derecho penal y la libertad de expresión en Ecuador.

El estudio se sustentó en una metodología documental y analítica, enfocada en examinar las tensiones existentes entre el avance de las tecnologías basadas en inteligencia artificial y la capacidad del ordenamiento jurídico ecuatoriano para responder frente a nuevas formas de afectación de bienes jurídicos protegidos.

Para el desarrollo de la investigación se efectuó una revisión documental y doctrinal de fuentes normativas, jurisprudenciales y bibliográficas relacionadas con la

inteligencia artificial, particularmente de la Constitución de la República del Ecuador [CRE] (2008), el Código Orgánico Integral Penal [COIP] (2025), la Ley Orgánica de Protección de Datos Personales [LOPDP] (2021) y la Ley Orgánica Electoral y de Organizaciones Políticas de la República del Ecuador, Código de la Democracia (2009); a su vez, se incorporó doctrina especializada relacionada con la inteligencia artificial, plataformas digitales, regulación tecnológica, desinformación automatizada y libertad de expresión.

De igual manera se aplicó una técnica de análisis jurídico interpretativo y comparado, mediante la cual se examinaron criterios doctrinarios, estándares interamericanos y problemáticas vinculadas con suplantación de identidad, *deepfakes*, fraude informático y tratamiento automatizado de datos.

3. Resultados

3.1 Inteligencia Artificial y Transformación del Ecosistema Comunicacional

El desarrollo de la inteligencia artificial ha modificado significativamente las dinámicas de producción, circulación y consumo de información en entornos digitales; en ese contexto las tecnologías basadas en aprendizaje automático y modelos generativos permiten automatizar procesos de creación de contenidos mediante sistemas capaces de producir textos, imágenes, audios y videos sintéticos con altos niveles de realismo.

Este escenario ha ampliado las capacidades de interacción digital, pero también ha incrementado los riesgos asociados con manipulación informativa, difusión

masiva de contenidos falsificados y utilización estratégica de herramientas automatizadas en plataformas digitales (Chesney & Citron, 2019).

La expansión de plataformas digitales y sistemas algorítmicos ha transformado el ecosistema comunicacional contemporáneo; de acuerdo con Van Dijck (2016), las plataformas operan como estructuras capaces de organizar la circulación de contenidos mediante mecanismos automatizados de visibilidad, segmentación y recomendación, influyendo directamente en la formación de opinión pública. En este contexto, la inteligencia artificial no solo facilita la automatización de contenidos, sino también la personalización de mensajes y la amplificación acelerada de información a gran escala.

Uno de los fenómenos más relevantes derivados del uso de inteligencia artificial generativa corresponde a los *deepfakes*, entendidos como contenidos audiovisuales manipulados digitalmente mediante algoritmos capaces de simular características físicas, voces o comportamientos humanos con apariencia verosímil; en ese contexto, según Chesney y Citron (2019), estas tecnologías plantean riesgos asociados con desinformación, afectación reputacional, fraude y vulneración de derechos personalísimos, especialmente cuando son utilizadas para inducir error o generar escenarios ficticios difíciles de distinguir de contenidos auténticos.

Al respecto, la automatización de perfiles digitales y la utilización de sistemas coordinados de difusión permiten incrementar artificialmente la circulación de determinados

mensajes en redes sociales y plataformas digitales; al respecto Woolley y Howard (2018) sostienen que la propagación automatizada de contenidos puede influir en procesos comunicacionales y debates públicos mediante estrategias de manipulación digital dirigidas a maximizar alcance, viralización e impacto emocional; es decir estas dinámicas adquieren especial relevancia en contextos políticos y electorales, donde la circulación masiva de información alterada o engañosa puede afectar la confianza pública y la integridad del debate democrático.

3.2 Conductas Ilícitas Emergentes Mediadas por Inteligencia Artificial

3.2.1 Suplantación de Identidad Digital. El uso de inteligencia artificial generativa ha facilitado la creación de contenidos audiovisuales sintéticos capaces de reproducir rasgos faciales, voces, expresiones y comportamientos humanos con altos niveles de precisión; por tal motivo, estas tecnologías, conocidas comúnmente como *deepfakes*, utilizan modelos de aprendizaje profundo para manipular o generar imágenes, audios y videos aparentemente auténticos, dificultando la diferenciación entre contenidos reales y material artificialmente alterado (Chesney & Citron, 2019).

La utilización de *deepfakes* plantea riesgos relevantes para la protección de derechos personalísimos, especialmente en supuestos relacionados con suplantación de identidad digital, afectación reputacional, vulneración de intimidad y difusión no consentida de contenidos manipulados; en línea con esto, la capacidad de simular declaraciones, comportamientos o imágenes de una persona

mediante herramientas automatizadas incrementa el potencial de engaño y amplificación masiva de contenidos falsificados en plataformas digitales, generando impactos tanto individuales como colectivos.

Desde una perspectiva jurídica, la utilización de contenidos sintéticos no constituye por sí misma una conducta necesariamente ilícita, en ese sentido, la valoración jurídica depende del contexto de utilización, la finalidad perseguida y la existencia de una afectación concreta a bienes jurídicos protegidos, además, en determinados casos, los *deepfakes* pueden emplearse con fines artísticos, satíricos o recreativos amparados por la libertad de expresión; sin embargo, su utilización para inducir error, afectar derechos personalísimos o ejecutar mecanismos de fraude digital puede generar consecuencias jurídicas relevantes.

En el ordenamiento jurídico ecuatoriano no existe actualmente un tipo penal específico que sancione de manera autónoma la creación o difusión de *deepfakes*. No obstante, ello no impide que determinadas conductas realizadas mediante estas tecnologías puedan ser analizadas a través de figuras penales ya previstas en el COIP (2025), especialmente cuando concurran elementos relacionados con fraude, afectación patrimonial, vulneración de la intimidad, uso indebido de datos personales o difusión de información de circulación restringida; así pues, la relevancia jurídico penal de los *deepfakes* debe determinarse a partir de las circunstancias concretas del caso y de la eventual afectación a bienes jurídicos protegidos.

La facilidad para producir contenidos sintéticos realistas incrementa los desafíos asociados con verificación

de información, confianza pública y protección de la identidad digital, particularmente en escenarios de alta viralización o circulación masiva de contenidos alterados.

Como efecto de lo anterior, el análisis jurídico de los *deepfakes* exige evitar respuestas basadas exclusivamente en expansión punitiva; por otro lado, la intervención penal únicamente resulta legítima cuando la utilización de estas tecnologías produce una afectación concreta y jurídicamente relevante sobre bienes protegidos por el ordenamiento jurídico, respetando los principios de legalidad, proporcionalidad y mínima intervención penal

3.2.2 Fraude Informático Automatizado. La incorporación de sistemas de inteligencia artificial en entornos digitales ha incrementado la sofisticación de mecanismos de fraude ejecutados mediante automatización tecnológica, particularmente en escenarios de interacción digital masiva y comunicación automatizada (Woolley & Howard, 2018).

Es decir, estas herramientas capaces de generar mensajes personalizados, replicar identidades digitales, producir voces sintéticas o automatizar interacciones comunicacionales han ampliado las posibilidades de ejecución de engaños dirigidos a obtener beneficios patrimoniales ilícitos mediante medios digitales; asimismo, la utilización de inteligencia artificial permite además incrementar el alcance, velocidad y capacidad de segmentación de estas prácticas mediante procesos automatizados de difusión e interacción algorítmica (Van Dijck, 2016).

Entre las modalidades más relevantes se encuentran los sistemas automatizados de *phishing*, la clonación de voz mediante inteligencia artificial, la simulación de identidades digitales y la utilización de contenidos sintéticos destinados a inducir error en contextos financieros, comerciales o personales; en ese sentido dichas herramientas, permiten recrear comunicaciones aparentemente auténticas mediante correos electrónicos, mensajes instantáneos, llamadas automatizadas o contenidos audiovisuales manipulados, generando escenarios de engaño complejos debido al alto nivel de verosimilitud que pueden alcanzar los sistemas generativos contemporáneos (Chesney & Citron, 2019).

Desde una perspectiva jurídico penal, el análisis de estas conductas exige diferenciar el uso legítimo de herramientas tecnológicas de aquellas prácticas orientadas a producir un perjuicio patrimonial mediante engaño; por consiguiente, la inteligencia artificial constituye un instrumento tecnológico que, por sí mismo, carece de relevancia penal autónoma; sin embargo, su utilización para ejecutar mecanismos fraudulentos puede permitir la adecuación de determinadas conductas a tipos penales ya previstos en el ordenamiento jurídico ecuatoriano; en este contexto, el elemento central del análisis no radica en la existencia de automatización tecnológica, sino en la concurrencia de engaño, error inducido, beneficio ilícito y afectación patrimonial.

De conformidad con lo que establece el COIP (2025), la estafa consiste en "la simulación de hechos falsos o deformación u ocultamiento de hechos verdaderos con la finalidad de obtener un beneficio patrimonial ilegítimo en

perjuicio de otra persona" (art. 186). Así pues, determinadas prácticas ejecutadas mediante inteligencia artificial podrían subsumirse parcialmente en esta figura cuando la automatización tecnológica sea utilizada como mecanismo para inducir error y provocar desplazamiento patrimonial.

Al respecto, el COIP (2025) incorpora disposiciones que permiten abordar determinadas conductas asociadas al uso indebido de inteligencia artificial en entornos digitales, entre ellas la estafa (art. 186), la revelación ilegal de bases de datos (art. 229), el acceso no consentido a un sistema informático (art. 230) y otras conductas relacionadas con la revelación de secretos o información personal de terceros. Estas figuras resultan relevantes para analizar posibles afectaciones al patrimonio, la privacidad, la seguridad informática y la protección de datos personales derivadas de prácticas digitales mediadas por inteligencia artificial.

No obstante, la utilización de inteligencia artificial en contextos fraudulentos también plantea desafíos relevantes en materia probatoria y atribución de responsabilidad; en ese sentido la automatización de interacciones digitales y la utilización de sistemas capaces de replicar patrones comunicacionales humanos dificultan la identificación de autores materiales, la trazabilidad de contenidos y la determinación de niveles de participación en entornos digitales complejos; estas dificultades se incrementan cuando intervienen plataformas transnacionales, mecanismos automatizados de difusión o herramientas de anonimización tecnológica.

Desde una perspectiva garantista, el análisis jurídico de estas conductas exige evitar interpretaciones

expansivas incompatibles con el principio de legalidad penal; en ese sentido, la mera utilización de herramientas automatizadas o sistemas de inteligencia artificial no justifica automáticamente intervención punitiva; es por ello que la respuesta penal únicamente resulta legítima cuando exista una afectación concreta a bienes jurídicos protegidos y concurran los elementos objetivos y subjetivos exigidos por el tipo penal correspondiente; como resultado, la utilización de inteligencia artificial debe analizarse como un posible medio de comisión de conductas ilícitas y no como una categoría autónoma de criminalidad desvinculada de los principios tradicionales del derecho penal.

3.2.3 Manipulación Informativa Automatizada y Desinformación Digital. La expansión de sistemas de inteligencia artificial aplicados a entornos digitales ha incrementado la capacidad de producción, segmentación y difusión automatizada de contenidos informativos a gran escala; es decir éstas herramientas capaces de generar textos, imágenes, audios y videos sintéticos permiten automatizar procesos de comunicación digital mediante mecanismos de personalización algorítmica y amplificación masiva de mensajes, modificando significativamente las dinámicas contemporáneas de circulación de información (Van Dijck, 2016). Por ende, la utilización de inteligencia artificial para producir o difundir contenidos engañosos plantea desafíos relevantes para la integridad del debate público, la confianza informativa y la protección de derechos fundamentales.

La desinformación digital constituye un fenómeno complejo que no puede reducirse únicamente a la existencia

Recibido: 17-04-2026
Aceptado: 29-05-2026
Publicado: 20-06-2026

de información falsa; en ese sentido, Wardle y Derakhshan (2017) sostienen que los procesos contemporáneos de desinformación incluyen mecanismos organizados de manipulación comunicacional orientados a influir en percepciones públicas mediante circulación estratégica de contenidos engañosos, descontextualizados o emocionalmente polarizantes; en ese sentido, la inteligencia artificial ha ampliado estas capacidades mediante automatización de perfiles digitales, generación masiva de contenidos sintéticos y utilización de sistemas capaces de segmentar audiencias específicas con altos niveles de precisión.

Entre las principales prácticas asociadas con manipulación informativa automatizada se encuentran la utilización coordinada de *bots*, generación masiva de contenidos falsificados, difusión automatizada de información manipulada y empleo de sistemas de inteligencia artificial para alterar artificialmente tendencias, interacciones o percepciones públicas en plataformas digitales. Al respecto Woolley y Howard (2018) advierten que, estos mecanismos de propaganda computacional pueden influir significativamente en procesos políticos, sociales y electorales mediante estrategias destinadas a maximizar viralización, impacto emocional y alcance comunicacional.

No obstante, desde una perspectiva jurídica y constitucional, el análisis de estos fenómenos exige especial cautela para evitar respuestas incompatibles con la libertad de expresión; en ese sentido, la circulación de información falsa o controversial no constituye automáticamente una conducta penalmente relevante.

Al respecto, la Corte IDH (2004) ha sostenido que la libertad de expresión protege no solo informaciones consideradas favorables o inofensivas, sino también aquellas que pueden resultar críticas, perturbadoras o incómodas para el Estado o determinados sectores sociales. Es decir, la intervención estatal frente a fenómenos de desinformación debe observar criterios estrictos de necesidad, proporcionalidad y legalidad.

Por otro lado, en el ordenamiento jurídico ecuatoriano no existe actualmente un tipo penal específico dirigido a sancionar la desinformación digital o la manipulación informativa automatizada. Esta ausencia normativa resulta relevante debido a que una regulación excesivamente amplia podría generar riesgos de censura indirecta o restricciones desproporcionadas al ejercicio de la libertad de expresión.

Por consiguiente, la CRE (2008) reconoce el derecho a buscar, recibir, intercambiar, producir y difundir información por cualquier medio, así como garantías relacionadas con libertad de pensamiento, opinión y comunicación.

Sin embargo, determinadas prácticas vinculadas con utilización de inteligencia artificial podrían adquirir relevancia jurídica cuando produzcan afectaciones concretas sobre bienes jurídicos protegidos; esto podría ocurrir, por ejemplo, en casos relacionados con fraude, vulneración de datos personales, suplantación de identidad digital, difusión no consentida de contenidos íntimos o utilización coordinada de mecanismos engañosos destinados a alterar ilícitamente procesos institucionales o electorales; en tales supuestos, el análisis jurídico no debe centrarse

exclusivamente en la falsedad del contenido difundido, sino en la existencia de daño jurídicamente relevante, dolo, afectación concreta y adecuación típica conforme a los principios del derecho penal.

De igual manera, el Código de la Democracia (2012) incorpora disposiciones orientadas a garantizar la transparencia, equidad y legitimidad de los procesos electorales; su pertinencia en el presente estudio radica en que las tecnologías de inteligencia artificial pueden facilitar la difusión automatizada de contenidos engañosos, la utilización coordinada de *bots* y la manipulación informativa masiva durante campañas electorales, situaciones que podrían afectar la formación libre de la voluntad política de los electores y la integridad del debate democrático.

Por consiguiente, la manipulación informativa automatizada representa uno de los principales desafíos contemporáneos para los sistemas democráticos y jurídicos digitales. Sin embargo, la complejidad del fenómeno exige respuestas normativas proporcionales y compatibles con derechos fundamentales, evitando recurrir a fórmulas expansivas de criminalización que puedan afectar indebidamente la libertad de expresión y el principio de legalidad penal.

3.3 Desafíos para el Derecho Penal Ecuatoriano

3.3.1 Principio de Legalidad y Adecuación Típica.

El surgimiento de conductas mediadas por inteligencia artificial plantea desafíos relevantes para el derecho penal contemporáneo, particularmente en relación con la capacidad de los ordenamientos jurídicos para responder frente a nuevas formas de afectación digital sin vulnerar las

garantías propias del Estado constitucional de derechos; en este contexto, el análisis jurídico de fenómenos asociados con inteligencia artificial exige observar estrictamente los principios de legalidad, tipicidad y mínima intervención penal, evitando interpretaciones expansivas incompatibles con las garantías fundamentales reconocidas en sistemas democráticos.

El principio de legalidad constituye uno de los límites esenciales del poder punitivo estatal y exige que toda conducta sancionada penalmente se encuentre previamente descrita de manera clara, expresa y taxativa en la ley penal. En el ordenamiento jurídico ecuatoriano, este principio se encuentra reconocido en el artículo 76 numeral 3 de la Constitución de la República del Ecuador (2008), así como en el artículo 5 del COIP (2025), el cual establece que no existe infracción penal, pena ni proceso penal sin ley previa; en efecto, desde la dogmática penal, este principio cumple una función garantista destinada a impedir la arbitrariedad estatal y limitar la expansión descontrolada del derecho penal (Roxin, 1997).

De esta manera, la aparición de tecnologías capaces de generar contenidos sintéticos, automatizar interacciones digitales o replicar identidades virtuales no implica automáticamente la existencia de nuevas categorías autónomas de criminalidad; es decir la inteligencia artificial constituye una herramienta tecnológica cuyo análisis jurídico debe realizarse a partir de conductas concretas y de la eventual afectación a bienes jurídicos protegidos por el ordenamiento penal. Como señala Mir Puig (2015), el derecho penal no sanciona tecnologías o instrumentos en

pp. 94-133
Recibido: 17-04-2026
Aceptado: 29-05-2026
Publicado: 20-06-2026

sí mismos, sino comportamientos humanos que lesionan o ponen en peligro bienes jurídicos relevantes conforme a los principios de lesividad y responsabilidad.

Desde esta perspectiva, el principal desafío jurídico consiste en determinar si determinadas prácticas mediadas por inteligencia artificial pueden subsumirse válidamente en tipos penales ya existentes sin recurrir a interpretaciones analógicas prohibidas en materia penal; lo cual, resulta especialmente relevante frente a fenómenos como *deepfakes*, automatización fraudulenta, manipulación digital o utilización indebida de datos personales, donde frecuentemente existe una tensión entre innovación tecnológica y capacidad regulatoria del derecho vigente.

El COIP (2025) incorpora diversas figuras que podrían resultar aplicables frente a determinadas conductas ejecutadas mediante inteligencia artificial, particularmente aquellas relacionadas con estafa (art. 186), revelación ilegal de bases de datos (art. 229), acceso no consentido a un sistema informático (art. 230) y revelación de secretos o información personal de terceros; sin embargo, la adecuación típica de estas conductas exige verificar rigurosamente la concurrencia de elementos objetivos y subjetivos del tipo penal correspondiente, especialmente en relación con engaño, dolo, afectación concreta y resultado jurídicamente relevante.

Asimismo, el análisis de adecuación típica debe considerar las dificultades interpretativas derivadas de entornos digitales complejos, donde pueden intervenir múltiples actores, plataformas tecnológicas y sistemas automatizados de difusión; estas circunstancias generan

desafíos relevantes relacionados con autoría, participación, imputación objetiva y atribución de responsabilidad penal, especialmente cuando las conductas son ejecutadas mediante herramientas algorítmicas capaces de operar de manera parcialmente automatizada.

Desde una perspectiva constitucional, cualquier intento de expansión del derecho penal frente a fenómenos asociados con inteligencia artificial debe interpretarse de manera restrictiva y compatible con el principio de proporcionalidad; al respecto la Corte Interamericana de Derechos Humanos, ha sostenido que las restricciones estatales que puedan afectar libertad de expresión o circulación de información deben satisfacer criterios estrictos de legalidad, finalidad legítima y necesidad en una sociedad democrática (Corte IDH, 2004).

Siguiendo esta línea, la respuesta penal frente a prácticas digitales mediadas por inteligencia artificial únicamente resulta legítima cuando exista una afectación concreta a bienes jurídicos protegidos y concurran de manera clara los presupuestos exigidos por el ordenamiento penal vigente.

Bajo este prisma, los desafíos derivados del uso de inteligencia artificial no justifican la flexibilización de garantías penales fundamentales; por el contrario, la complejidad tecnológica contemporánea exige reforzar criterios de interpretación restrictiva, legalidad y protección de derechos fundamentales, evitando respuestas basadas exclusivamente en expansión punitiva frente a fenómenos digitales emergentes

3.3.2 Protección de Bienes Jurídicos en Entornos Digitales. La utilización de inteligencia artificial en entornos digitales plantea desafíos relevantes para la protección de diversos bienes jurídicos tradicionalmente tutelados por el derecho constitucional y penal; en ese sentido, la capacidad de los sistemas automatizados para generar contenidos sintéticos, procesar grandes volúmenes de información, replicar identidades digitales y amplificar mensajes mediante mecanismos algorítmicos puede producir afectaciones sobre derechos relacionados con intimidad, patrimonio, autodeterminación informativa y seguridad de las comunicaciones digitales.

Desde la teoría del bien jurídico, la intervención penal únicamente resulta legítima cuando existe una afectación concreta o una puesta en peligro relevante respecto de intereses fundamentales protegidos por el ordenamiento jurídico (Roxin, 1997).

Por consiguiente, el análisis de conductas mediadas por inteligencia artificial no debe centrarse exclusivamente en la tecnología utilizada, sino en la existencia de lesión o riesgo jurídicamente relevante respecto de bienes protegidos por la Constitución y la ley; esta línea argumentativa, según Mir Puig (2015), el derecho penal no tiene como finalidad sancionar innovaciones tecnológicas o riesgos abstractos, sino proteger bienes jurídicos frente a afectaciones graves incompatibles con la convivencia social y el orden constitucional.

Uno de los principales ámbitos de protección corresponde a la intimidad y a los datos personales; en ese contexto, la Constitución de la República del Ecuador (2008), “reconoce el derecho a la protección de datos de

carácter personal, incluyendo acceso, decisión y control sobreinformación de esta naturaleza, así como garantías relacionadas con honra, imagen e intimidad personal y familiar (art. 66). en concordancia con ello, la LOPDP (2021) establece principios de juridicidad, finalidad, proporcionalidad y tratamiento legítimo aplicables al procesamiento de datos personales en entornos digitales, incluyendo categorías sensibles vinculadas con información biométrica, imagen, voz y perfiles digitales.

La relevancia de estos derechos adquiere especial importancia frente a sistemas de inteligencia artificial generativa que utilizan datos personales para entrenamiento algorítmico, reconocimiento de patrones o creación de contenidos sintéticos; considerando lo expuesto, la utilización no consentida de imágenes, registros biométricos, voces o información personal para generación de *deepfakes* o simulaciones digitales puede afectar derechos relacionados con identidad, autodeterminación informativa e intimidad, particularmente cuando dichos contenidos son difundidos masivamente o utilizados con fines engañosos.

Asimismo, determinadas prácticas mediadas por inteligencia artificial pueden afectar bienes jurídicos patrimoniales mediante mecanismos de fraude automatizado orientados a inducir error en contextos financieros, comerciales o personales.

En tales casos, la afectación penalmente relevante no se configura por la mera automatización tecnológica, sino por la concurrencia de engaño, desplazamiento patrimonial y beneficio ilícito conforme a los elementos exigidos por el COIP (2025). Esta distinción resulta fundamental para evitar

pp. 94-133
Recibido: 17-04-2026
Aceptado: 29-05-2026
Publicado: 20-06-2026

interpretaciones expansivas incompatibles con el principio de legalidad penal.

Otro ámbito particularmente sensible corresponde a la protección de los procesos comunicacionales y democráticos frente a mecanismos coordinados de manipulación digital masiva. La utilización de perfiles automatizados, sistemas algorítmicos de amplificación y contenidos sintéticos puede influir artificialmente en dinámicas de circulación informativa y deliberación pública en plataformas digitales.

En ese sentido Woolley y Howard (2018) sostienen que las estrategias de propaganda computacional permiten amplificar determinados mensajes mediante automatización y segmentación digital masiva, generando riesgos para transparencia informativa y confianza pública en contextos electorales y comunicacionales.

No obstante, la protección de estos bienes jurídicos debe interpretarse de manera compatible con los estándares constitucionales e interamericanos de libertad de expresión; al decir, la Corte Interamericana de Derechos Humanos ha señalado que las restricciones estatales sobre circulación de información deben satisfacer criterios estrictos de legalidad, necesidad y proporcionalidad, evitando mecanismos ambiguos o excesivamente amplios que puedan producir censura indirecta o afectar el debate democrático (Corte IDH, 2004).

Esto implica que, la protección jurídica frente a fenómenos digitales no puede fundamentarse exclusivamente en la existencia de información falsa, controversial o socialmente cuestionada, sino en la concurrencia de una afectación concreta y jurídicamente

relevante sobre bienes protegidos por el ordenamiento jurídico.

Desde esta perspectiva, los desafíos derivados del uso de inteligencia artificial exigen respuestas jurídicas integrales orientadas a proteger derechos fundamentales sin recurrir automáticamente a expansión punitiva; es decir, la complejidad de los entornos digitales contemporáneos requiere articular mecanismos constitucionales, regulatorios, tecnológicos y eventualmente penales bajo criterios de proporcionalidad, legalidad y mínima intervención estatal.

3.3.3 Riesgos de Expansión Punitiva y Sobrecriminalización. La aparición de nuevas formas de interacción digital mediadas por inteligencia artificial ha generado presiones crecientes para ampliar los mecanismos de control estatal frente a fenómenos asociados con desinformación, automatización comunicacional y manipulación digital; al respecto, desde una perspectiva constitucional y garantista, la incorporación de respuestas penales frente a estas problemáticas exige especial cautela para evitar procesos de expansión punitiva incompatibles con los principios fundamentales que limitan el ejercicio del poder sancionador del Estado.

La expansión del derecho penal hacia ámbitos caracterizados por innovación tecnológica y circulación masiva de información plantea riesgos relevantes de indeterminación normativa y afectación desproporcionada de derechos fundamentales, en ese contexto, Silva Sánchez (2001) advierte que las sociedades contemporáneas tienden a responder frente a nuevos riesgos sociales mediante ampliación constante del derecho penal, fenómeno que

puede conducir a debilitamiento progresivo de garantías tradicionales como legalidad, taxatividad y mínima intervención; es decir, en contextos digitales, esta tendencia adquiere especial relevancia debido a la dificultad para delimitar con precisión qué conductas justifican intervención penal y cuáles permanecen amparadas por derechos fundamentales.

Uno de los principales riesgos asociados con criminalización expansiva en entornos digitales consiste en la utilización de categorías ambiguas o excesivamente amplias vinculadas con «desinformación», «manipulación digital» o «contenidos falsos»; mismo que carecen de delimitación normativa estricta, pueden generar escenarios de inseguridad jurídica incompatibles con el principio de legalidad penal reconocido en el artículo 76 numeral 3 de la CRE (2008) y en el artículo 5 del COIP (2025). Desde la dogmática penal, el principio de taxatividad exige que las conductas sancionadas penalmente se encuentren descritas de manera clara, previa y estricta, evitando fórmulas abiertas que permitan interpretaciones arbitrarias (Roxin, 1997).

De igual modo, la expansión punitiva frente a fenómenos de circulación informativa puede afectar de manera significativa la libertad de expresión y el debate democrático, en ese sentido la Corte Interamericana de Derechos Humanos ha sostenido reiteradamente que la libertad de expresión constituye uno de los fundamentos esenciales de las sociedades democráticas y que las restricciones estatales deben interpretarse restrictivamente para evitar mecanismos directos o indirectos de censura (Corte IDH, 2004).

Derivado de lo anterior, la existencia de información errónea, controversial o socialmente cuestionada no justifica automáticamente intervención penal, especialmente cuando no existe una afectación concreta a bienes jurídicos protegidos.

En el contexto digital contemporáneo, la tensión entre combate a la desinformación y protección de libertad de expresión adquiere especial complejidad debido a la velocidad de circulación informativa y a la capacidad de amplificación algorítmica de plataformas digitales; al respecto, advierte Balkin (2018), las respuestas regulatorias orientadas a controlar contenidos digitales deben evitar convertirse en mecanismos de vigilancia o censura incompatibles con los principios democráticos y constitucionales.

Desde esta perspectiva, la inteligencia artificial no debe ser concebida como una categoría autónoma de peligrosidad que justifique flexibilización de garantías penales. El análisis jurídico de conductas mediadas por sistemas automatizados debe continuar sujeto a los principios tradicionales del derecho penal, particularmente legalidad, lesividad, culpabilidad y proporcionalidad. La mera utilización de herramientas tecnológicas avanzadas no convierte automáticamente una conducta en penalmente relevante, ni habilita, interpretaciones extensivas incompatibles con un Estado constitucional de derechos.

En el caso ecuatoriano, la ausencia de tipos penales específicos dirigidos a sancionar fenómenos de desinformación digital o manipulación automatizada obliga a interpretar restrictivamente las disposiciones existentes, evitando recurrir a analogías penales prohibidas

o construcciones doctrinales expansivas destinadas a suplir vacíos regulatorios mediante ampliación artificial de tipos penales vigentes. Esta exigencia resulta particularmente importante en escenarios donde pueden verse comprometidos derechos relacionados con libertad de expresión, acceso a la información y participación democrática.

De ahí que, los desafíos derivados del uso de inteligencia artificial en entornos digitales requieren respuestas jurídicas compatibles con los estándares constitucionales e interamericanos de protección de derechos fundamentales.

La complejidad de estos fenómenos exige fortalecer mecanismos regulatorios, educativos, tecnológicos y de protección de datos personales antes que recurrir automáticamente a expansión del derecho penal como respuesta principal frente a riesgos digitales emergentes. Desde una perspectiva garantista, la intervención penal únicamente resulta legítima cuando exista una afectación concreta a bienes jurídicos protegidos y concurran de manera estricta los elementos exigidos por el ordenamiento jurídico vigente.

3.4 Libertad de Expresión, Proporcionalidad y Regulación Digital

La expansión de tecnologías basadas en inteligencia artificial ha intensificado los debates contemporáneos relacionados con libertad de expresión, regulación digital y límites legítimos de intervención estatal en entornos comunicacionales automatizados. La capacidad de producir y difundir contenidos sintéticos mediante sistemas algorítmicos plantea desafíos relevantes para los Estados democráticos,

particularmente frente a fenómenos asociados con desinformación, manipulación digital y circulación masiva de contenidos alterados. No obstante, cualquier respuesta normativa dirigida a enfrentar estas problemáticas debe interpretarse de manera compatible con las garantías constitucionales y los estándares internacionales de protección de la libertad de expresión.

La libertad de expresión constituye uno de los pilares fundamentales de las sociedades democráticas y representa una condición esencial para el ejercicio de otros derechos fundamentales. La Constitución de la República del Ecuador (2008), en su artículo 66 numeral 6, reconoce el derecho a opinar y expresar libremente el pensamiento en todas sus formas y manifestaciones, mientras que el artículo 16 garantiza el derecho a una comunicación libre, intercultural, incluyente, diversa y participativa en todos los ámbitos de interacción social.

Estos estándares se complementan con lo dispuesto en el artículo 13 de la Convención Americana sobre Derechos Humanos, el cual protege la libertad de buscar, recibir y difundir informaciones e ideas de toda índole por cualquier medio; igualmente, la Corte Interamericana de Derechos Humanos ha señalado reiteradamente que la libertad de expresión posee una dimensión individual y colectiva indispensable para el funcionamiento del sistema democrático (Corte IDH, 2004).

Bajo esta perspectiva, la protección convencional no se limita únicamente a expresiones consideradas verdaderas, aceptadas o socialmente favorables, sino que también comprende aquellas opiniones o informaciones que puedan resultar críticas, incómodas o perturbadoras para

determinados sectores sociales o instituciones estatales.

Desde este enfoque, las restricciones estatales sobre circulación de información deben satisfacer estrictamente los principios de legalidad, finalidad legítima, necesidad y proporcionalidad. En entornos digitales, estos principios adquieren especial relevancia debido a la capacidad de plataformas tecnológicas y sistemas algorítmicos para influir en dinámicas de circulación informativa y moderación de contenidos.

Balkin (2018) sostiene que las sociedades contemporáneas enfrentan crecientes tensiones entre gobernanza digital, libertad de expresión y poder de intermediación tecnológica ejercido por plataformas privadas capaces de determinar visibilidad, permanencia o difusión de contenidos en espacios digitales.

La utilización de inteligencia artificial para moderación automatizada de información también plantea riesgos relacionados con opacidad algorítmica, censura indirecta y restricciones desproporcionadas sobre el debate público digital; desde esta perspectiva, los fenómenos de desinformación y manipulación digital no pueden abordarse exclusivamente mediante respuestas punitivas o mecanismos amplios de restricción de contenidos.

La utilización de categorías ambiguas como «información falsa» o «desinformación» puede generar riesgos de arbitrariedad e inseguridad jurídica cuando no existen criterios normativos claros para delimitar conductas jurídicamente relevantes. Como advierte la Relatoría Especial para la Libertad de Expresión de la Comisión Interamericana de Derechos Humanos (2019), los Estados deben evitar

regulaciones excesivamente amplias que permitan sancionar expresiones legítimas bajo argumentos vinculados con combate a la desinformación o protección del orden público.

Por otra parte, el principio de proporcionalidad exige que toda medida restrictiva relacionada con circulación de contenidos digitales sea idónea, necesaria y estrictamente proporcional respecto del fin legítimo perseguido; esto implica que la intervención penal debe constituir un mecanismo excepcional y subsidiario frente a fenómenos de manipulación digital o automatización comunicacional.

De lo expuesto se deduce que, antes de recurrir a respuestas sancionatorias, los Estados deben priorizar estrategias orientadas a alfabetización digital, transparencia algorítmica, protección de datos personales, fortalecimiento institucional y mecanismos regulatorios compatibles con estándares democráticos y constitucionales.

En el contexto ecuatoriano, estos desafíos adquieren especial importancia debido a la coexistencia de garantías constitucionales robustas de libertad de expresión y crecientes riesgos asociados con utilización indebida de inteligencia artificial en entornos digitales. En ese sentido, la ausencia de regulación específica sobre inteligencia artificial obliga a interpretar las disposiciones existentes de manera compatible con principios constitucionales e interamericanos, evitando construcciones expansivas que permitan restringir indebidamente el debate público o criminalizar expresiones protegidas por el ordenamiento jurídico.

Se colige que, la regulación de fenómenos asociados con inteligencia artificial y comunicación digital exige mantener un equilibrio entre protección de derechos fundamentales y preservación de libertades democráticas.

Recibido: 17-04-2026
Aceptado: 29-05-2026
Publicado: 20-06-2026

3.5 Protección de Datos Personales y Responsabilidad Digital

El desarrollo de sistemas de inteligencia artificial aplicados a entornos digitales ha incrementado significativamente los desafíos relacionados con protección de datos personales, tratamiento automatizado de información y responsabilidad derivada del uso de tecnologías algorítmicas.

La capacidad de los sistemas de inteligencia artificial para recopilar, procesar, almacenar y analizar grandes volúmenes de datos mediante mecanismos automatizados genera riesgos relevantes para derechos vinculados con intimidad, autodeterminación informativa y control sobre información personal, especialmente cuando intervienen categorías sensibles de datos como registros biométricos, imágenes, patrones de voz o perfiles digitales.

La CRE (2008), en su artículo 66 numeral 19, reconoce el derecho a la protección de datos de carácter personal, incluyendo acceso, decisión, control y disposición sobre información personal, así como garantías relacionadas con honra, imagen e intimidad.

Este reconocimiento constitucional adquiere especial relevancia frente a tecnologías capaces de utilizar datos personales para entrenamiento algorítmico, reconocimiento automatizado de patrones o generación de contenidos sintéticos mediante inteligencia artificial.

En concordancia con el texto constitucional, la LOPDP (2021) establece principios orientados a garantizar tratamiento legítimo, proporcional y seguro de datos personales en entornos físicos y digitales. Entre estos

principios destacan juridicidad, lealtad, transparencia, finalidad, pertinencia y minimización de datos, los cuales resultan especialmente relevantes frente a sistemas automatizados capaces de procesar información personal a gran escala.

La LOPDP también reconoce protección reforzada respecto de categorías especiales de datos personales, incluyendo información biométrica y datos sensibles cuyo tratamiento indebido puede afectar gravemente derechos fundamentales.

La utilización de inteligencia artificial generativa y sistemas automatizados de reconocimiento digital plantea desafíos relevantes en relación con tratamiento de datos biométricos e información personal sensible. Herramientas capaces de replicar rostros, voces, expresiones faciales o patrones conductuales mediante algoritmos de aprendizaje profundo requieren frecuentemente utilización masiva de imágenes, registros audiovisuales o información personal obtenida de plataformas digitales y entornos virtuales.

En este contexto, la creación de *deepfakes* o simulaciones digitales mediante utilización no consentida de datos biométricos puede generar afectaciones relacionadas con identidad digital, intimidad y autodeterminación informativa. De la misma manera, el Reglamento General a la Ley Orgánica de Protección de Datos Personales (2023) incorpora criterios relacionados con gestión de riesgos, responsabilidad proactiva y medidas de seguridad aplicables al tratamiento automatizado de datos personales.

Estas disposiciones adquieren relevancia frente a sistemas de inteligencia artificial capaces de adoptar

decisiones automatizadas o procesar información mediante mecanismos algorítmicos complejos, especialmente cuando existe posibilidad de afectación significativa sobre derechos y libertades fundamentales de las personas titulares de datos. Desde esta perspectiva, la responsabilidad digital derivada del uso de inteligencia artificial no puede limitarse exclusivamente a análisis de responsabilidad penal individual.

La complejidad de los ecosistemas digitales contemporáneos exige considerar también obligaciones relacionadas con tratamiento legítimo de datos, deberes de seguridad, transparencia algorítmica y prevención de riesgos tecnológicos por parte de desarrolladores, operadores de plataformas y responsables del tratamiento de información personal. Como resultado de ello, la utilización de inteligencia artificial implica deberes jurídicos que trascienden el ámbito estrictamente sancionatorio y se proyectan hacia mecanismos integrales de protección de derechos fundamentales en entornos digitales.

No obstante, la regulación de tecnologías basadas en inteligencia artificial también debe interpretarse de manera compatible con principios constitucionales de proporcionalidad e innovación tecnológica; en ese sentido, la protección de datos personales no implica prohibición absoluta del desarrollo de sistemas automatizados o herramientas digitales avanzadas, sino exigencia de mecanismos adecuados de tratamiento legítimo, consentimiento informado, transparencia y prevención de afectaciones indebidas sobre derechos fundamentales.

En esta misma línea, la regulación jurídica debe orientarse a garantizar equilibrio entre innovación

tecnológica, desarrollo digital y protección efectiva de derechos fundamentales. A raíz de esto, los desafíos contemporáneos relacionados con inteligencia artificial y protección de datos personales exigen fortalecer marcos regulatorios orientados a garantizar responsabilidad digital, transparencia algorítmica y tutela efectiva de derechos fundamentales en entornos digitales.

La complejidad tecnológica contemporánea requiere respuestas jurídicas integrales que articulen garantías constitucionales, regulación especializada y mecanismos efectivos de control sobre tratamiento automatizado.

4. Discusión

Los resultados obtenidos evidencian que el desarrollo de la inteligencia artificial generativa ha transformado significativamente las dinámicas tradicionales de comunicación digital, ampliando las capacidades de producción automatizada de contenidos y facilitando nuevas formas de interacción en entornos virtuales; sin embargo, estas mismas herramientas también han incrementado los riesgos asociados con manipulación informativa, suplantación de identidad digital, utilización indebida de datos personales y difusión masiva de contenidos sintéticos potencialmente lesivos para bienes jurídicos protegidos.

En este contexto, el análisis realizado demuestra que el ordenamiento jurídico ecuatoriano permite abordar parcialmente determinadas conductas derivadas del uso indebido de inteligencia artificial mediante figuras penales ya existentes, especialmente aquellas relacionadas con fraude informático, acceso no consentido a sistemas digitales, estafa y vulneración de información personal previstas en el COIP (2025). No obstante, también se

identificaron limitaciones normativas frente a fenómenos emergentes, especialmente la difusión de *deepfakes* que afectan la reputación o la identidad digital sin encajar claramente en tipos penales vigentes, así como los mecanismos automatizados de manipulación informativa y desinformación digital que, aun cuando pueden generar efectos socialmente relevantes, no siempre presentan una adecuación típica suficiente dentro del régimen penal ecuatoriano.

Estas dificultades regulatorias derivan, en gran medida, de la acelerada evolución tecnológica y de la complejidad que supone atribuir responsabilidad jurídica frente a conductas mediadas por sistemas automatizados de inteligencia artificial. A diferencia de las formas tradicionales de criminalidad informática, muchas prácticas digitales contemporáneas operan mediante mecanismos descentralizados, automatizados y transnacionales, dificultando la identificación de sujetos responsables, la delimitación típica de las conductas y la determinación del alcance del daño producido.

Por otra parte, la investigación permitió advertir que la expansión de respuestas sancionatorias frente a fenómenos de desinformación automatizada puede generar riesgos de sobrecriminalización incompatibles con el principio de mínima intervención penal y con las garantías derivadas de la libertad de expresión. En sociedades democráticas, la circulación de información incluso cuando resulte controversial, errónea o ideológicamente confrontativa permanece protegida por estándares constitucionales e interamericanos que limitan la posibilidad

de establecer restricciones estatales desproporcionadas.

En este sentido, la Corte Interamericana de Derechos Humanos ha sostenido reiteradamente que la libertad de expresión constituye uno de los pilares esenciales de las sociedades democráticas y que cualquier restricción estatal debe superar estrictos parámetros de necesidad, proporcionalidad y legalidad; se deduce que, la utilización del derecho penal como mecanismo de respuesta frente a fenómenos vinculados con inteligencia artificial y desinformación digital debe aplicarse de manera excepcional, evitando fórmulas normativas ambiguas o expansivas que puedan producir efectos inhibitorios sobre el debate público y la circulación legítima de información.

De igual manera, el estudio evidencia que la protección de datos personales adquiere especial relevancia frente al uso de sistemas automatizados capaces de recopilar, procesar y perfilar información digital a gran escala. Aunque la LOPDP (2021) incorpora principios orientados a garantizar el tratamiento legítimo de datos, persisten desafíos regulatorios relacionados con utilización de datos biométricos, procesamiento automatizado de información y opacidad algorítmica en plataformas digitales.

Desde una perspectiva jurídico-penal, los hallazgos permiten concluir que no toda conducta socialmente problemática derivada del uso de inteligencia artificial debe traducirse automáticamente en nuevas formas de criminalización. La expansión indiscriminada del derecho penal frente a fenómenos tecnológicos emergentes podría generar respuestas simbólicas carentes de eficacia práctica y potencialmente incompatibles con el principio

pp. 94-133
Recibido: 17-04-2026
Aceptado: 29-05-2026
Publicado: 20-06-2026

de legalidad penal. En consecuencia, cualquier política regulatoria orientada a enfrentar riesgos digitales asociados con inteligencia artificial debe articular mecanismos proporcionales, técnicamente precisos y compatibles con la protección simultánea de bienes jurídicos y derechos fundamentales.

Finalmente, el análisis desarrollado permite sostener que los desafíos jurídicos derivados de la inteligencia artificial no pueden abordarse exclusivamente desde una perspectiva sancionatoria. La complejidad del ecosistema digital contemporáneo exige estrategias integrales que incorporen mecanismos de alfabetización digital, transparencia algorítmica, responsabilidad tecnológica y fortalecimiento institucional, evitando respuestas regulatorias simplificadas frente a fenómenos tecnológicos caracterizados por su constante transformación y dinamismo.

5. Conclusiones

La investigación evidenció que cualquier respuesta estatal orientada a regular conductas derivadas del uso indebido de inteligencia artificial debe observar estrictamente los principios de legalidad penal, proporcionalidad y mínima intervención. La expansión indiscriminada del derecho penal frente a fenómenos tecnológicos emergentes podría generar riesgos de sobre criminalización incompatibles con la libertad de expresión y con los estándares constitucionales e interamericanos aplicables en sociedades democráticas.

De igual manera, se constató que la protección de datos personales constituye uno de los principales

desafíos regulatorios en escenarios de automatización digital, particularmente frente al uso masivo de sistemas de procesamiento y perfilamiento de información mediante inteligencia artificial. Aunque la LOPDP (2021) incorpora mecanismos relevantes de tutela jurídica, todavía existen desafíos relacionados con transparencia algorítmica, utilización de datos biométricos y responsabilidad digital en plataformas tecnológicas.

Finalmente, se concluye que los desafíos jurídicos derivados de la inteligencia artificial no pueden abordarse exclusivamente mediante respuestas sancionatorias.

6. Referencias

Balkin, J. M. (2018). Free speech in the algorithmic society: Big data, private governance, and new school speech regulation. *UC Davis Law Review, 51*(3), 1149-1210. https://lawreview.law.ucdavis.edu/sites/g/files/dgvnsk15026/files/media/documents/51-3_Balkin.pdf

Castells, M. (2009). *Comunicación y poder*. Alianza Editorial.

Chesney, R. & Citron, D. K. (2019). Deep fakes: A looming challenge for privacy, democracy, and national security. *California Law Review, 107*(6), 1753-1820. https://doi.org/10.15779/Z38RV0D15J

Código Orgánico Integral Penal [COIP] de 2025. Registro Oficial Suplemento 180 de 10 de febrero de 2014. Última Reforma: Edición Constitucional del Registro Oficial 96, 03-X-2025

Comisión Interamericana de Derechos Humanos, Relatoría Especial para la Libertad de Expresión. (2019). *Guía para garantizar la libertad de expresión frente a la desinformación deliberada en contextos electorales*. Comisión Interamericana de Derechos Humanos. https://www.oas.org/es/cidh/expresion/publicaciones/Guia_Desinformacion_VF.pdf

Constitución de la República del Ecuador de 2008. Registro Oficial No. 449 de 20 de octubre de 2008.

Convención Americana sobre Derechos Humanos de 1969. San José de Costa Rica el 22 de noviembre de 1969. https://www.oas.org/es/cidh/mandato/documentos-basicos/convencion-americana-derechos-humanos.pdf

Corte Interamericana de Derechos Humanos [Corte IDH]. (2004). *Caso Herrera Ulloa vs. Costa Rica. Sentencia de 2*

de julio de 2004. https://www.corteidh.or.cr/docs/casos/articulos/seriec_107_esp.pdf

Ley Orgánica de Protección de Datos Personales [LOPDP]. (2021). Registro Oficial Suplemento 459 de 26 de mayo de 2021.

Ley Orgánica Electoral y de Organizaciones Políticas de la República del Ecuador, Código de la Democracia de 2009. Registro Oficial Suplemento 578 de 27 de abril de 2009.

Mir Puig, S. (2015). *Derecho penal: Parte general* (10.ª ed.). Reppertor.

Reglamento General a la Ley Orgánica de Protección de Datos Personales. (2023). Decreto Ejecutivo No. 904. Registro Oficial Suplemento 435 de 14 de noviembre de 2023.

Roxin, C. (1997). *Derecho penal: Parte general*. Civitas.

Silva Sánchez, J. M. (2001). *La expansión del derecho penal: Aspectos de la política criminal en las sociedades postindustriales* (2.ª ed.). Civitas.

Van Dijck, J. (2016). *La cultura de la conectividad: Una historia crítica de las redes sociales*. Siglo XXI Editores.

Wardle, C. & Derakhshan, H. (2017). *Information disorder: Toward an interdisciplinary framework for research and policy making*. Council of Europe.

Woolley, S. C. & Howard, P. N. (2018). *Computational propaganda: Political parties, politicians, and political manipulation on social media*. Oxford University Press.