La dosificación penal en los delitos de difamación e injuria cometidos mediante inteligencia artificial

Sorelis Marín Aponte[1]
Universidad José María Vargas
sorelismarin@gmail.com
ORCID 0009-0009-7477-7756

Jenny M. González Franquis[2]
Universidad Santa María
jmgfranquis@hotmail.com
ORCID 0009-0000-4664-8935

Juan Pablo Torres[3]
Universidad Santa María
jptorres@oediciones.com
ORCID 0009-0007-9616-0605

Tomás A. Pérez Ruíz[4]
Universidad Santa María
psjuridicofinanciero@gmail.com
ORCID 0009-0004-7185-5468

[1] Sorelis Marín Aponte (Venezuela), abogada (U. José María Vargas); especialista y magíster en Derecho Penal (USM/UNE), doctoranda en Derecho Internacional Privado y en Educación (UPEL). Abogada en libre ejercicio; experiencia en civil, LOPNNA y violencia de género.

[2] Jenny Mercedes González Franquis (Venezuela), abogada (USM); jueza provisoria del Juzgado Sexto de Municipio del AMC y docente de Derecho Procesal Civil (UCSAR). Doctora en Derecho Constitucional y Penitenciario (CIU) y en Ciencias de la Educación (ULAC); máster en Derecho Constitucional (UNED).

[3] Juan Pablo Torres Delgado (Venezuela), abogado (USM); Doctorado en Dercho Constitucional (CIU), postdoctorado en Seguridad y Defensa de la Nación (IAEDEN), Ciencias Jurídicas y Políticas (UNE), y en Ciencias de la Educación (UBA).

[4] Tomás A. Pérez Ruíz (Venezuela), abogado y licenciado en Contaduría Pública (USM); doctor en Derecho Constitucional y Penitenciario (CIU) y máster en Derecho Constitucional (UNED). Profesor (USM/IUPOL-UNES); experiencia en litigación penal y asesoría financiera y tributaria.

La dosificación penal en los delitos de difamación e injuria cometidos mediante inteligencia artificial

Resumen

El artículo examina la dosificación penal de los delitos de difamación e injuria cuando la inteligencia artificial (IA) funge como medio comisivo. Sostiene tres premisas: la dignidad como bien jurídico tutelado que integra honor y reputación; el carácter humano del autor, siendo la IA un instrumento; y la incorporación del "daño causado" —moral/reputacional, integridad física/psicológica, económico-patrimonial, profesional/laboral, familiar/social y privacidad— como eje para graduar la pena. Propone una matriz Daño × Gravedad, agravantes y atenuantes específicos del uso de IA (p. ej., verosimilitud sintética, automatización/escala, opacidad técnica, alcance/irreversibilidad), y un procedimiento operativo (verificación estructural, pericia técnica en IA, preservación de evidencia y motivación estandarizada) con formatos tipo para fiscalías y juzgados. En diálogo con marcos internacionales (AI Act de la UE, Convenio Marco del Consejo de Europa, OCDE/UNESCO) y el contexto venezolano, argumenta que una dosificación proporcional y explícita refuerza la tutela judicial efectiva, eleva el efecto disuasorio y corrige la subvaloración punitiva de estos ilícitos cuando median tecnologías de IA. Concluye con recomendaciones de implementación inmediata, capacitación y cooperación para evidencia electrónica transfronteriza.

Palabras clave: inteligencia artificial (IA); difamación; injuria; dosificación penal; matriz daño–gravedad; honor y reputación; deepfakes (verosimilitud sintética); prueba digital y cadena de custodia; tutela judicial efectiva; Venezuela

Penalty Dosimetry for Defamation and Insult Offenses Committed Through Artificial Intelligence

Abstract

This article analyzes sentencing (dosimetry) for defamation and insult when artificial intelligence (AI) is used as the means of commission. It advances three pillars: human dignity as the protected legal interest encompassing honor and reputation; the human perpetrator, with AI as an instrument; and the express integration of "harm caused"—moral/reputational, physical/psychological, economic-property, professional/employment, family/social, and privacy—as the backbone for penalty calibration. The paper introduces a Harm × Severity matrix, AI-specific aggravating and mitigating factors (e.g., synthetic verisimilitude, automation/scale, technical opacity, reach/irreversibility), and an operational workflow (structural verification, AI forensics, evidence preservation, and standardized reasoning) supported by ready-to-use forms for prosecutors and courts. Anchored in international frameworks (EU AI Act, Council of Europe Framework Convention, OECD/UNESCO) and the Venezuelan legal setting, it contends that proportional, explicitly reasoned sentencing strengthens effective judicial protection, enhances general deterrence, and addresses the current under-penalization of AI-mediated offenses. The article closes with immediate implementation steps, capacity building, and cross-border electronic-evidence cooperation.

Keywords: artificial intelligence (AI); defamation; insult (injuria); sentencing (penal dosimetry); harm–severity matrix; honor and reputation; deepfakes (synthetic verisimilitude); digital evidence and chain of custody; effective judicial protection; Venezuela

Introducción

En Venezuela, una persona identificada como “A” alberga odio, discriminación o una rivalidad laboral contra otra persona identificada como “B”, quien —a fuerza de trabajo y educación— ha alcanzado un alto cargo y prestigio social, lo que lo hace más vulnerable al escarnio público. “A” hace un uso indebido de la inteligencia artificial (IA) y construye un video de “B” que lo muestra en una situación indecorosa, contraria a la moral y a las buenas costumbres, socialmente reprochable y sancionada por la opinión pública. Luego lo difunde a través de los diversos medios disponibles en cualquier teléfono celular: Instagram, Facebook, YouTube, X y WhatsApp. El resultado es un rechazo social de tal magnitud que “B” no puede sostener su empleo ni su prestigio; en consecuencia, pierde su sustento económico y ve afectadas sus finanzas personales, el orden familiar, la educación de sus hijos, el estándar de vida, la medicina que sostiene la salud de sus padres y su círculo social y laboral, que se reduce hasta cerrarle puertas en su entorno profesional.

Cuando “B” decide ejercer acciones legales contra “A”, su denuncia se tipifica como “delito menor”, pues nuestro Código Penal establece penas de uno a tres años para la difamación y de seis meses a un año para la injuria. Así, “A” —quien ha dañado una vida profesional, familiar y social— no es sancionado con prisión efectiva. Tampoco resulta aplicable la “confiscación y supresión de los impresos, dibujos y demás objetos que hayan servido para cometer el delito” (artículo 448, eiusdem), ya que los medios utilizados no permiten cuantificar las veces que el video ha sido reproducido y reenviado, volviendo el daño prácticamente irreversible. Nuestra legislación, por tanto, luce limitada para alcanzar su finalidad.

En otras partes del mundo, “el bullying y el ciberbullying son causantes de más de 200 mil muertes en todo el planeta cada año, según un estudio de la Organización Mundial de la Salud y Bullying Sin Fronteras. Todos los días hay casos de suicidios, de homicidios y de tentativas de ambos. La salud y la seguridad mental y física de nuestros chicos está en peligro en todo el mundo; sin embargo, el *establishment* político internacional, la mayoría de los docentes y otros responsables de darnos una mano en la lucha contra el *bullying* y el *ciberbullying*, miran de soslayo, afirmó el Dr. Javier Miglino, experto en Derechos Humanos y Protección de la Niñez y voz autorizada mundial en la materia. Se trata de niños y adolescentes cuyas familias enfrentan un vacío irrecuperable; historias de suicidios provocados por acoso cibernético; difamaciones e injurias que, además, continúan replicándose en Internet día tras día.

Los ejemplos citados son historias reales, que cada día tocan a un amigo, familiar, cónyuge, hijo, sobrino, vecino o compañero de trabajo. Su gravedad obliga a alzar la voz, el silencio es parte del problema. Es necesaria la participación activa para aplicar el principio de progresividad de la ley, a fin de actualizar los elementos del delito, garantizar el debido proceso y la tutela judicial efectiva, y honrar así nuestra tan amada Justicia. En este artículo plasmamos reflexiones que parten de una serie de interrogantes cuyas respuestas exigirían no solo la reforma de las leyes sustantivas, sino también ajustes en las leyes adjetivas y en los procedimientos, en coherencia con la tutela judicial efectiva.

El principio de progresividad de la ley impone adaptar las normas a medida que la sociedad cambia, avanza y se desarrolla en los ámbitos social, económico, familiar, científico, profesional y laboral, entre otros. En ese marco, los autores de este artículo estamos convencidos de la necesidad social de sensibilizar a los legisladores para actualizar y reformar los parámetros que rigen los delitos penales cometidos con IA —en

particular, la difamación y la injuria—. Ello no será posible sin incluir dentro de su tipicidad el "daño causado" y sin visualizar el conjunto de variables que inciden en estos delitos, con la mira puesta en aportar criterios jurídicos que permitan determinar una dosimetría (o dosificación) penal adecuada a su fin último: prevenir y reprimir la violación de la norma, para hacer prevalecer su cumplimiento.

Abordamos, además, un glosario de términos poco o nunca utilizados en normativa positiva y proponemos varios enfoques con óptica jurídica para los delitos de difamación e injuria cuando el medio comisivo es la IA. Nos adentramos en bases conceptuales como "dignidad", "tutela", "bien jurídico tutelado" y "derechos humanos", teniendo como nexo el uso de la IA como instrumento delictivo, con el objeto de evidenciar la urgencia de adecuar el debido proceso y la tipificación a una dosificación penal sincerada con la realidad social.

La aceleración de la IA ha ampliado el repertorio y el alcance de conductas ilícitas mediadas por tecnología —desde fraudes con *deepfakes* y suplantaciones de identidad, hasta campañas coordinadas de desinformación—, tensionando los sistemas penales en todo el mundo. Informes recientes de autoridades europeas advierten que la IA está "turboimpulsando" al crimen organizado y complejizando la labor probatoria y de persecución penal (Europol, 2025; AP News, 2025; Reuters). Este panorama ha motivado respuestas regulatorias de gran escala, por ejemplo, la Unión Europea promulgó la AI Act (en vigor desde el 1 de agosto de 2024, con aplicación escalonada hasta 2027), que inaugura un régimen por niveles de riesgo para sistemas de IA —incluidos los modelos fundacionales— y prevé prohibiciones específicas con implicaciones para la trazabilidad y la prueba digital en procesos penales (Comisión Europea, 2024; White & Case, 2024) [Estrategia Digital]. En paralelo, el Convenio Marco sobre IA del Consejo de Europa —primer tratado internacional vinculante en la materia— exige

que todo el ciclo de vida de la IA sea compatible con los derechos humanos, la democracia y el Estado de Derecho, creando un lenguaje común con potencial de irradiación hacia legislaciones nacionales y comparadas (Consejo de Europa, 2024a; Consejo de Europa, 2024b; Covington, 2024) [Portal]. Estas iniciativas se apoyan en estándares previos, como los Principios de la OCDE sobre IA (2019, actualización 2024) y la Recomendación de la UNESCO sobre la Ética de la IA (2021/2024), que refuerzan valores de proporcionalidad, seguridad, responsabilidad y debido proceso aplicables al ámbito penal (OCDE, 2019/2024; UNESCO, 2021/2024) [oecd.ai].

Venezuela no es ajena a esta realidad. En el contexto nacional, la IA ya opera como medio para la comisión de conductas que van desde “bromas” y “chanzas” hasta difamación, injuria, calumnia, extorsión y estafa, con daños reputacionales, económicos y familiares potencialmente irreversibles. Sin embargo, estos hechos suelen tratarse como “delitos menores” bajo los rangos sancionatorios vigentes, lo que debilita su efecto disuasorio y la tutela efectiva de las víctimas. La base de esta investigación caracteriza la dosificación penal de delitos cometidos mediante IA como un tema emergente del derecho penal que reaviva debates sobre imputación, autoría, dolo/culpa y daño causado, y alerta sobre el riesgo de impunidad si el sistema no se adapta con criterios claros y proporcionales.

Aunque Estados Unidos y la Unión Europea avanzan en marcos de referencia, aún no existe un consenso internacional. Venezuela tiene, por ello, la oportunidad de actuar tempranamente y convertirse en referente regional si actualiza su enfoque de tipificación y dosificación penal frente a la IA.

Sobre esta base, la investigación propone interrogantes cuyas respuestas —aportadas por el lector— lo convierten en copartícipe de las premisas necesarias para “humanizar” el análisis de la pena, incorporando el daño efectivo (moral, físico,

económico, profesional, familiar y de privacidad) y su gravedad como elemento a considerar en la tipificación. Partimos de una arquitectura conceptual donde el sujeto activo es humano y la IA funge como instrumento del hecho punible. Esta aproximación busca alinear la respuesta penal con estándares globales de proporcionalidad y prevención del daño, al tiempo que atiende particularidades locales y vacíos normativos que hoy erosionan la prevención general y la protección de bienes jurídicos especialmente vulnerables en la era de la manipulación mediante inteligencia artificial.

Antecedentes Fácticos

Las leyes y la normativa sobre difamación e injuria existen desde hace varias décadas; sin embargo, el honor y la reputación como bienes protegidos se reconocen —desde la Edad Media— en estrecha relación con los derechos de la persona. A lo largo de distintos hitos históricos, su concepto y aplicación han evolucionado. En la actualidad, las penas oscilan entre multas administrativas y hasta tres años de prisión, e incluso hay jurisdicciones en las que estos asuntos no se tramitan como delitos penales, sino ante la jurisdicción civil. Así, los delitos de difamación, injuria y, en algunos casos, calumnia, reciben tratamientos punitivos diversos:

- **Perú:** la pena máxima por calumnia puede ser de hasta tres años de prisión.
- **Colombia:** por injuria puede imponerse prisión de dieciséis a cincuenta y cuatro meses y multa.
- **España:** las injurias leves pueden sancionarse con multas, mientras que las difundidas con publicidad pueden conllevar prisión o multas más elevadas. Su Código Penal, en el artículo 264, detalla penas para quienes difundan o utilicen información personal robada con el fin de dañar la reputa-

ción de la víctima.

- **Brasil:** las penas pueden ser de tres meses a un año de prisión, además de multas.
- **República Dominicana:** pueden imponerse de tres meses a un año de prisión y multas.
- **Venezuela:** las penas pueden incluir de uno a tres años de prisión y multas.
- **Europa:** el tratamiento varía según cada país. Se observan tendencias hacia la despenalización o el traslado al ámbito civil; en todo caso, rige la *Declaración Universal de Derechos Humanos*, cuyo artículo 12 reconoce la protección del honor y la reputación como derecho internacional.

Esta panorámica evidencia que, por la antigüedad de muchas leyes sustantivas, no se previeron los avances tecnológicos que hoy enfrentamos. La inteligencia artificial, bajo el principio de progresividad, obliga a adoptar planteamientos distintos a los que inspiraron las amplias exposiciones de motivos del pasado. A nuestro juicio, existen dos variables que deben incorporarse —en las distintas etapas del proceso— y que hoy no forman parte ni de la denuncia ni del trámite: la dignidad como principio y bien jurídico tutelado, y el daño causado como elemento del delito. En consecuencia, la pena debiera elevarse sustancialmente como medida de prevención, pues, conforme a máximas de experiencia, los delitos por estos conceptos bordean la impunidad: los procesos son tan exigentes y los resultados tan limitados que la víctima, con frecuencia, opta por la inacción u omite accionar ante violaciones constitucionales evidentes.

A medida que esa impunidad se afianza, la comisión de tales delitos se hace más común y repetida, con creciente osadía e, incluso, en abierto desafío a la norma y a las autoridades policiales, fiscales y judiciales.

Marco conceptual

Este marco fija un vocabulario común y los ejes analíticos para la dosificación penal cuando la IA interviene en la comisión de un delito. Enfatiza que la IA es medio y no autor, y que la graduación de la pena debe reflejar el daño efectivamente causado y la gravedad del resultado en la víctima.

Sujeto activo: El sujeto activo del delito es siempre una persona humana. Aun cuando se utilicen sistemas de IA, la responsabilidad recae en quien decide, configura o emplea ese recurso tecnológico para ejecutar la conducta típica. La IA, por sí misma, carece de capacidad jurídica para ser sujeto activo.

a) **IA como medio (instrumento):** La IA opera como medio o instrumento que potencia el alcance, la velocidad y la verosimilitud de la conducta. Por ejemplo: generación o difusión de contenidos manipulados). Para la dosificación interesa cómo se utilizó el medio: su función en la ejecución, el grado de intervención humana y el nexo causal con el resultado.

b) **Elemento subjetivo: dolo y culpa:**

- **Dolo:** deriva de la combinación sujeto–medio; exige la intención de emplear la IA para causar un daño (v. gr., estafa, difamación, injuria).
- **Culpa:** también puede configurarse (p. ej., falta de supervisión de representantes respecto de un menor que usa IA), lo que demanda criterios diferenciados de reproche y pena.

c) **Bienes jurídicos y daño:** La respuesta penal debe humanizar el análisis incorporando el daño real y sus dimensiones: moral, integridad física, estabilidad económica, esfera

profesional y/o familiar, privacidad, así como la gravedad del resultado. Estos elementos deben expresarse de manera explícita al graduar la pena.

d) **Relación con tipos penales relevantes:** Aunque el artículo prioriza los delitos de difamación e injuria por su frecuente subvaloración punitiva y el carácter, a veces, irreversible del daño, el enfoque es trasladable a otros tipos (p. ej., calumnia, estafa) cuando la IA se utiliza como medio.

e) **Finalidad preventiva y proporcionalidad:** La hipótesis central sostiene que una dosificación proporcional y visible a la gravedad del daño —cuando la IA es el medio delictivo— contribuye a la prevención general y reduce la incidencia de estas conductas.

f) **El bien jurídico tutelado:** En este apartado realizamos un análisis exegético del **bien jurídico tutelado** con el objeto de sensibilizar al lector sobre la necesidad de reforzar la respuesta frente a los delitos de difamación e injuria cometidos mediante IA como medio comisivo.

g) **Tutela:** Aunque la *tutela* es, en su origen, una institución del derecho civil que atribuye a una persona la guarda de los bienes o intereses de otra, en el ámbito penal —y para nuestros fines pedagógicos— el término designa la función protectora del Estado sobre los derechos fundamentales y constitucionales cuando resultan amenazados o vulnerados por particulares o por entes públicos, a través de personas o de sus acciones.

h) **Bien jurídico:** El bien jurídico es el interés o valor socialmente relevante que el ordenamiento reconoce y protege. Sus antecedentes se inscriben en la tradición jurídica europea y su consolidación moderna se afirma con la codificación liberal y, tras la Segunda Guerra Mundial, cobra nuevo impulso con los juicios de Núremberg, que fijaron garantías mínimas para todo proceso judicial frente a los excesos de regímenes totalitarios, y con la consagración de derechos

en constituciones como la Ley Fundamental de la República Federal de Alemania.

En esta línea, Rivera (2002) entiende por *constitucionalización del derecho* el proceso de incorporación a la ley suprema de normas que limitan el poder del Estado y establecen parámetros superiores —en especial sobre la ley procesal— para hacer efectivas las libertades y la tutela de los derechos de las personas. Así se reconocen, de manera intrínseca al ser humano, derechos como salud, propiedad, medio ambiente, honor, reputación y libertad, dotados de un interés tangible o intangible y protegidos como derechos fundamentales. Con el devenir social, este catálogo se ha complementado mediante la incorporación de nuevos derechos y el fortalecimiento de mecanismos para sancionar conductas que atenten contra ellos.

i) **Bien jurídico tutelado:** Conforme a lo expuesto, los bienes jurídicos tutelados se definen por el valor que las personas y la comunidad asignan a aquello que hace posible una vida digna; el Estado debe protegerlos frente a ataques provenientes de personas o entes, públicos o privados.

La Constitución y las normas que de ella derivan cumplen una función de garantía propia del Estado de Derecho, que exige identificar el bien jurídico tutelado lesionado y, en consecuencia, remite al principio de tipicidad. En el ámbito penal, esta identificación vincula los hechos con el derecho fundamental afectado y constituye el contenido básico de la tipicidad.

j) **Principio de legalidad:** El principio de legalidad impone que los poderes públicos actúen siempre sometidos a la ley preexistente y conforme a la Constitución, en su tiempo, espacio y modo. De él se derivan, entre otros, los postulados *nullum crimen, nulla poena sine lege* y la prohibición de la analogía en perjuicio del reo (*in malam partem*).

Ello exige que las conductas prohibidas estén descriptas previamente (tipificación) y que a cada descripción se asocie una sanción; de este modo se prohíbe crear, con posterioridad, normas o sanciones no previstas.

k) **Dignidad:** En los delitos de difamación e injuria, el bien jurídico tutelado es el honor y la reputación. No obstante, ambos se hallan íntima e intrínsecamente vinculados con la dignidad humana.

La dignidad es un concepto unificador que concentra aquello que confiere al ser humano su suprema valía interior y le permite definirse como sujeto autónomo, dotado de independencia. Supone relacionar a la persona con su identidad, igualdad y no discriminación; con su actividad económica, laboral y profesional; con su apariencia física y psíquica; con su propiedad, familia, libertad (incluida la libertad de expresión y la libertad sexual) y su religión. En consecuencia, el derecho constitucional —y las normas penales que de él derivan— incorpora la dignidad humana como valor innato y fundamento jurídico que justifica la protección del honor y la reputación.

l) **Fundamento constitucional venezolano:**
La **Constitución de la República Bolivariana de Venezuela** reconoce y protege la dignidad en múltiples disposiciones, entre ellas:

Artículo 3: *"El Estado tiene como fines esenciales la defensa y el desarrollo de la persona y el* ***respeto a su dignidad****, el ejercicio democrático de la voluntad popular, la construcción de una sociedad justa y amante de la paz, la promoción de la prosperidad y bienestar del pueblo y la garantía del cumplimiento de los principios, derechos y deberes reconocidos y consagrados en esta Constitución. La* ***educación*** *y el* ***trabajo*** *son los procesos fundamentales para alcanzar dichos fines".*

Artículo 46: *"[…]* ***Toda persona privada de libertad será tratada con el respeto debido a la dignidad*** *inherente al ser humano".*

Artículo 47: *"El hogar doméstico y todo recinto privado de persona son inviolables. No podrán ser allanados sino mediante orden judicial […]* ***respetando siempre la dignidad del ser humano****".*

Artículo 55: *"[…]* ***Los cuerpos de seguridad del Estado respetarán la dignidad y los derechos humanos*** *de todas las personas".*

Artículo 80: *"El Estado garantizará a los ancianos y ancianas el pleno ejercicio de sus derechos y garantías […]* ***respetar su dignidad humana****…".*

Artículo 81: *"Toda persona con discapacidad o necesidades especiales tiene derecho al ejercicio pleno y autónomo de sus capacidades […]* ***se le garantizará el respeto a su dignidad humana****…".*

Artículo 91: *"Todo trabajador o trabajadora tiene derecho a un salario suficiente que le permita* ***vivir con dignidad****…".*

Reconocimiento internacional

La protección de la dignidad cuenta, además, con amplio **respaldo internacional**:

- **Ley Fundamental de la República Federal de Alemania, art. 1:** *"La dignidad del hombre es intangible. Respetarla y protegerla es obligación de todo poder público".*
- **Declaración Universal de Derechos Humanos, art. 12:** *"Nadie será objeto de injerencias arbitrarias en su vida privada, su familia, su domicilio o su correspondencia, ni de ataques a su honra o a su reputación. Toda persona tiene derecho a la protección de la ley contra tales injerencias o ataques".*

En suma, la dignidad no es un concepto meramente metafísico: ha sido positivizada en sus diversas manifestaciones como derecho humano y derecho constitucional, y, por tanto, como bien jurídico tutelado cuya salvaguarda corresponde al Estado. Desde esta perspectiva, el honor y la reputación —lesionados por la difamación y la injuria, incluso cuando se valen de la IA como medio— se protegen precisamente porque su afectación compromete el núcleo de la dignidad de la persona.

Difamación e injuria: Asumiendo la dignidad como bien jurídico tutelado, y siendo consistente con que tanto la injuria como la difamación atentan contra el honor y la reputación —y, por ende, contra la propia dignidad de la persona—, cabe preguntarse: ¿por qué no otorgar a los delitos de difamación e injuria la misma protección que se reconoce a otros bienes jurídicos tutelados que integran la dignidad?

Para nosotros, la respuesta es evidente: debe dárseles la misma protección, tutela y debido proceso.

En otras palabras: ante la violación del derecho a la vida, el homicidio es atendido de inmediato por el Ministerio Público —incluso de oficio— y se inicia una investigación en la que se activan técnicas criminalísticas y se recaban pruebas para hacer justicia. Algo similar ocurre cuando se lesionan bienes jurídicos vinculados con la integridad física, la no discriminación, los delitos de odio, la libertad sexual o la violencia de género.

En cambio, cuando se vulneran el honor y la reputación —que igualmente forman parte del haz de derechos que conforman la dignidad—, el itinerario procesal suele ser distinto: con frecuencia requiere denuncia a instancia de parte, desplaza en gran medida el acervo probatorio hacia la víctima y, por regla general, se judicializa como delito menor. Este tratamiento desigual resulta especialmente problemático cuando la IA se ha utilizado como medio comisivo, pues multiplica el alcance, la persistencia y la irreversibilidad del daño.

Criterios operativos de dosificación: Como en toda norma penal, no basta con fijar un parámetro o banda de pena; es imprescindible delinear los criterios agravantes y atenuantes del delito. Para ello, deben considerarse factores que inciden de forma directa en la determinación y aplicabilidad de la sanción. En este sentido, presentamos el análisis que, a nuestro juicio, debe integrarse a los criterios operativos de dosificación.

El objetivo es la proporcionalidad: que la pena refleje el daño efectivo y la gravedad del resultado cuando la IA funge como medio comisivo, y que, a la vez, contribuya a la prevención de estas conductas, sin convertir esta propuesta en un cuadro sancionatorio cerrado o mecánico.

Verificación estructural del caso (véase Anexo 2):

Antes de encuadrar o tipificar el delito cometido mediante IA, es imprescindible identificar los elementos básicos del hecho punible. Ello exige:

1. **Autor humano e IA como medio.** Determinar que el sujeto activo es una persona y que la IA actúa exclusivamente como instrumento.

2. **Elemento subjetivo (dolo o culpa).** Verificar la intención de utilizar la IA para el ilícito o, en su caso, la culpa por falta de supervisión (p. ej., de representantes respecto de un menor).

3. **Nexo de causalidad.** Establecer el vínculo entre la conducta del autor y el resultado dañoso, precisando que, en el caso analizado, el uso de la IA es el medio que conecta al autor con el dolo o la culpa para causar el daño.

Es en este último punto donde ingresamos en un terreno aún poco explorado por la literatura penal. Para identificar y valorar el nexo causal, debemos dimensionar el alcance de las

consecuencias de la acción del autor según los instrumentos tecnológicos empleados. Aquí aparece un abanico de variables —criterios, puntos de vista, tipos de tecnología, niveles de habilidad técnica— que carecen de límites previamente establecidos. Esta indeterminación dificulta el encuadre preciso del elemento que, a nuestro juicio, debe incorporarse expresamente a la tipificación del delito: *el daño causado*.

El daño causado

Para integrar un estándar objetivo en la estructura del caso, proponemos evaluar el daño causado a partir de dimensiones verificables. Estas dimensiones deben constar en autos con evidencia suficiente y reflejarse en la dosificación de la pena cuando la IA es el medio comisivo:

a) Entorno de la víctima.

Contexto personal y de vulnerabilidad (exposición pública, dependencia económica, condición de especial protección, etc.).

Evidencia sugerida: testimonios, informes psicosociales, documentación de rol público.

b) Daño moral y reputacional.

Afectación al honor, estima social y confianza pública.

Evidencia: peritajes de reputación digital, hemerografía, rastreos en buscadores y redes.

c) Integridad física / violencia, acoso y hostigamiento.

Amenazas, seguimiento, doxxing, ciberacoso u otras conductas que incrementen el riesgo físico o psicológico.

Evidencia: denuncias, capturas de mensajes, informes médicos/psicológicos.

d) Afectación económica/patrimonial.

Pérdida de ingresos, contratos, clientela o incremento de costos (defensa, contención reputacional).

Evidencia: cartas de despido o rescisión, estados de cuenta, facturas y peritajes contables.

e) Esfera profesional y de empleo.

Desprestigio laboral, inhabilitación de hecho, barreras para recolocación.

Evidencia: comunicaciones empresariales, rechazos documentados, referencias laborales.

f) Ámbito familiar y social.

Deterioro de relaciones, afectación a hijos/as (p. ej., bullying escolar), aislamiento comunitario.

Evidencia: informes escolares, intervención de trabajo social, testimonios.

g) Privacidad e intimidad.

Difusión de datos sensibles, imágenes íntimas o información privada.

Evidencia: peritajes sobre contenido, metadatos, registros de publicación.

h) Sexo y edad de la víctima (especial protección).

Mayor reproche cuando concurren minoridad, adultez mayor, embarazo, discapacidad u otras condiciones protegidas.

Evidencia: documentos de identidad, certificados médicos.

i) Medio utilizado: propaganda, divulgación o publicidad.

Naturaleza del canal y calidad verosímil del material (p. ej., deepfake), uso de automatización o bots.

Evidencia: peritaje forense digital, registros de plataforma, análisis técnico del contenido.

j) Alcance del medio utilizado.

Difusión efectiva: visualizaciones, compartidos, embeds, cobertura en prensa o en múltiples plataformas.

Evidencia: analíticas, logs, capturas verificadas, monitoreo de redes.

k) Irreversibilidad.

Persistencia del daño por replicabilidad, caches, indexación y dificultad de retirada.

Evidencia: reportes de desindexación/retirada y su (in)eficacia, rastreo de copias residuales.

> **Regla de aplicación:** cada dimensión debe valorarse y motivarse expresamente en la sentencia (baja / media / alta / muy alta), de modo que la suma cualitativa de impactos guíe la proporcionalidad de la pena y las medidas de reparación.

Agravantes específicos del uso de IA

Como hemos señalado, son múltiples los factores que deben considerarse al fijar la banda de pena y sus atenuantes/agravantes. La dificultad aumenta cuando la difamación o la injuria se cometen utilizando IA como medio, pues los canales de difusión tienen un alcance potencialmente ilimitado y la reversión del daño es, en muchos casos, imposible. Por ejemplo, en WhatsApp no existe una trazabilidad pública que permita

determinar cuántas veces fue reenviado un mismo mensaje, lo que frustra intentos de rectificación pública o indemnización capaces de restituir realmente la situación anterior: quien recibió el contenido lesivo no necesariamente verá la información que lo desmiente o repara.

Lejos de un tono fatalista, el análisis se agrava al constatar que la IA permite verosimilitud sintética del contenido: clonación de voz, texto altamente persuasivo, imágenes o videos *deepfake* que sitúan a personas en contextos inexistentes. Su retransmisión puede inducir, incluso ante autoridades policiales, fiscales o judiciales, falsos supuestos de hecho y/o de derecho, afectando la percepción sobre la víctima. Hoy, los autores pueden acceder —con barreras de entrada mínimas y simples tutoriales— a herramientas para producir videos, voces, hackeos, noticias falsas, manipulación algorítmica, usurpación de identidad, automatización con bots, uso de VPN/TOR, identidades ficticias, borrado de metadatos, etc.; prácticas asociadas a estafas y daños patrimoniales que con frecuencia quedan impunes. A ello se suma la exposición temprana de niñas, niños y adolescentes a entornos de IA, lo que incrementa el deber de protección estatal.

En un plano más organizado, redes delictivas que incurren en asociación para delinquir operan con fines económicos o desestabilizadores, incluso con efectos en la política de un país. Las víctimas sufren amenazas, acoso, y daños familiares, laborales y sociales, cuando no atentados contra su propia vida.

Componente de género y grupos especialmente protegidos

En difamación e injuria con IA, merecen un tratamiento específico los casos con componente de género y los menores de edad:

- **Mujeres como víctimas.** La imputación falsa suele centrarse en la vida privada, la moral sexual o la honradez, ámbitos íntimamente vinculados a la dignidad. Estas campañas constituyen violencia psicológica, con riesgos de ansiedad, depresión y aislamiento social, y pueden utilizarse para desacreditar a mujeres en posiciones de poder o en su entorno laboral, perpetuando estereotipos y discriminación. Cuando la difamación se realiza con menosprecio al género, debe considerarse agravante, aun cuando su prueba resulte compleja.

- **Asimetrías normativas.** A diferencia del marco especializado en violencia contra la mujer, no existe siempre un régimen específico equivalente para hombres víctimas de violencia psicológica cometida mediante IA; esta asimetría puede traducirse en tutelas y procedimientos menos idóneos, en tensión con principios constitucionales de igualdad y no discriminación.

- **Niñas, niños y adolescentes.** La normativa especial de protección brinda herramientas procesales y sancionatorias que pueden suplir la ausencia de una ley específica sobre IA, pero la hiperexposición digital y la replicabilidad del daño justifican agravación y medidas reforzadas de reparación.

Para el resto de la sociedad, las normas penales y procesales generales siguen aplicándose, pero la dinámica técnica de la IA y la asimetría de información entre víctima y autor aconsejan criterios agravados cuando concurran factores como los que siguen a continuación.

Criterios sugeridos de agravación (cuando media IA)

1. **Alcance y replicabilidad.** Difusión masiva, multicanal y transnacional; facilidad de **reenvío** y **reindexación**.
2. **Irreversibilidad.** Persistencia del contenido (copias, *cachés*, *mirrors*) e ineficacia práctica de las medidas de retirada o desindexación.
3. **Verosimilitud sintética.** Uso de deepfakes, clonación de voz u otros medios que elevan la credibilidad del contenido.
4. **Anonimato y ocultación.** Empleo de VPN/TOR, identidades ficticias, borrado de metadatos o plataformas opacas.
5. **Automatización y amplificación.** Uso de bots, granjas de clics o redes programadas para maximizar alcance e impacto.
6. **Coordinación delictiva.** Participación de grupos o asociación para delinquir, con fines de lucro, extorsión o desestabilización.
7. **Reiteración y persecución continuada.** Conducta sostenida en el tiempo, acoso, hostigamiento y re-victimización.
8. **Vulnerabilidad de la víctima.** Condición de menor edad, persona mayor, discapacidad, embarazo u otras circunstancias de especial protección.
9. **Componente de género o discriminación.** Menosprecio por sexo, género, orientación, etnia, etc., como móvil o contexto.
10. **Impacto económico y profesional.** Pérdida de empleo, contratos, clientela o imposibilidad de recolocación.
11. **Resistencia a la reparación.** Dificultad objetiva para rectificar, compensar o restaurar la situación previa.

> **Aplicación práctica.** Cada criterio debe ser expresamente motivado en la sentencia judicial (bajo/medio/alto/muy alto) y puede operar acumulativamente para desplazar la pena dentro de la banda o, cuando el legislador lo disponga, activar marcos agravados.

Factores atenuantes

Manteniendo el hilo de lo ya expuesto, y en opinión de los autores, cuando los delitos de difamación e injuria se cometen utilizando IA y redes digitales, resulta extraordinariamente complejo alcanzar una reparación justa del daño mediante retractación, retiro diligente del contenido o disculpa pública. Salvo supuestos muy excepcionales —en los que sea posible identificar plena y exclusivamente al público destinatario—, es casi imposible lograr una limpieza reputacional efectiva o la reparación del daño psicológico causado.

Aun así, corresponde considerar atenuantes bajo condiciones estrictas de verificación, entre ellas:

1. **Retractación eficaz, pronta y equivalente.**

   Valora la tempestividad (antes de la intervención fiscal/judicial), el alcance equivalente al daño (mismas plataformas/canales y visibilidad comparable), la claridad del desmentido y su permanencia.

   Condición probatoria: constancias técnicas de publicación y métricas de alcance.

2. **Retirada y desindexación diligentes.**

   Acciones verificables para eliminar copias, solicitar bajas y desindexación en buscadores y plataformas, así como avisos a replicadores relevantes.

   Condición probatoria: tickets/recibos de retiro, respuestas de plataformas, reportes forenses.

3. **Cooperación técnica con la investigación.**

   Preservación y entrega de evidencia digital (metadatos, *logs*, historiales), identificación de coautores y cesación voluntaria de la conducta.

   Condición probatoria: actas de entrega y peritajes.

4. **Reparación integral y medidas restaurativas.**

   Indemnización proporcional (daño patrimonial y gastos de defensa/atención psicológica), servicios profesionales de gestión reputacional y disculpa pública cuando sea idónea.

   Condición probatoria: acuerdos, transferencias, facturas y constancias de cumplimiento.

5. **Mitigación del riesgo de repetición.**

   Implementación de **controles** (cierres de cuentas, políticas internas, *compliance* y alfabetización digital), con verificación de su efectividad.

   Condición probatoria: protocolos, certificaciones, auditorías.

6. **Arrepentimiento activo**

   Cese espontáneo de la conducta y acciones de contención antes de cualquier requerimiento formal.

   Condición probatoria: trazabilidad temporal y evidencias de retiro voluntario.

7. **Difusión acotada y escasa replicabilidad.**

   Alcance limitado (círculo reducido, sin multicanal ni viralización), ausencia de verosimilitud sintética (contenido burdo que reduce credibilidad).

   Condición probatoria: analíticas y peritaje sobre calidad/verosimilitud.

8. Participación secundaria o marginal.

Rol **accesorio** respecto del plan principal (p. ej., retransmisor no coordinado), sin beneficio económico ni incentivo propio.

Condición probatoria: análisis de cadenas de difusión y ausencia de coordinación.

9. Capacidad disminuida o discapacidad

Situaciones que afecten la imputabilidad o la capacidad de comprensión y autodeterminación, debidamente acreditadas.

Condición probatoria: peritajes médicos/psicológicos.

10. Menores de edad y deber de supervisión.

Cuando el autor sea menor, rige su régimen especial; si lo cometido deriva de culpa por falta de supervisión de representantes, la valoración del reproche será diferenciada. A la inversa, la diligencia parental demostrable puede operar como atenuante en sede correspondiente.

Condición probatoria: constancias de controles parentales y actuaciones preventivas.

Regla de aplicación

- Los atenuantes solo proceden si su eficacia es demostrable y motivada en la resolución judicial.
- Su concurrencia no neutraliza la irreversibilidad ni el alcance del daño cuando éstos resulten altos.
- Deben ponderarse conjuntamente con los agravantes específicos de IA y con las dimensiones del daño ya definidas, para ubicar la sanción dentro de la banda de manera proporcional.

Justificación de la presente investigación

El auge de la IA como medio comisivo ha amplificado el daño moral, profesional, económico y familiar en víctimas de difamación e injuria, mientras que el marco sancionatorio vigente en Venezuela tiende a tratar estos hechos como "delitos menores", con escaso efecto disuasorio y serias dificultades para la reparación efectiva. Esta investigación propone humanizar la dosificación penal incorporando el daño real y la gravedad del resultado en la decisión judicial, y sostiene que un aumento proporcional de las penas —cuando la IA es el medio delictivo— puede tener impacto preventivo relevante, fortaleciendo la tutela judicial efectiva y la confianza en el sistema.

En el plano global, diversas autoridades y organismos han advertido que la IA potencia modalidades delictivas —*deepfakes*, suplantaciones, fraude, desinformación—, generando retos probatorios y de persecución penal que exigen actualizaciones normativas y criterios operativos claros (v. gr., Europol). En este contexto, el presente trabajo aporta: (i) la delimitación del bien jurídico tutelado y el rol de la IA como medio y no autor; (ii) una matriz de daño para orientar la graduación de la pena; (iii) agravantes y atenuantes específicos del uso de IA; y (iv) lineamientos para motivar decisiones proporcionales y coherentes con el principio de progresividad.

Delimitación

El estudio se circunscribe a Venezuela y, dentro del catálogo penal, prioriza los delitos de difamación e injuria cometidos mediante IA debido a su frecuencia, a su subvaloración punitiva y al potencial de irreversibilidad del daño. La IA se concibe como medio/instrumento, no como autor; el sujeto activo es siempre humano. La dosificación debe atender al elemento subjetivo (*dolo o culpa*), a las dimensiones del daño y a la gravedad del resultado.

Este trabajo no fija un cuadro sancionatorio cerrado, y contiene premisas orientadoras para eventuales reformas.

Análisis crítico

Se identifican tres brechas principales:

a) **Brecha sancionatoria.**

La calificación frecuente como “delitos menores” y los rangos punitivos vigentes no internalizan la magnitud ni la persistencia en línea del daño causado con IA (p. ej., dificultad de desindexación, huella digital perdurable), lo que erosiona el efecto disuasorio.

b) **Brecha probatoria y de trazabilidad**

La ausencia de pautas claras para valorar la verosimilitud sintética (deepfakes, clonación de voz), la automatización de la difusión y la opacidad técnica vuelven desigual la graduación de la pena. Este trabajo propone tratarlas como agravantes operativas y aportar criterios de evaluación técnica y su consideración en la motivación judicial.

c) **Brecha de armonización.**

Venezuela carece de un marco que dialogue con los estándares internacionales emergentes (UE/Consejo de Europa/ OCDE/UNESCO). Adoptar premisas de dosificación compatibles con dichos estándares permitiría desarrollar el marco regulatorio y dar mayor protección de los bienes jurídicos (v. gr., documentos del Consejo de Europa disponibles en rm.coe.int).

Pertinencia

La investigación es pertinente y oportuna porque provee un andamiaje operativo —premisas de dosificación y matriz daño × gravedad— que permite cerrar el déficit de proporcionalidad en los delitos contra el honor cometidos mediante IA. Con ello, se alinea el sistema penal venezolano con principios internacionales y con la realidad técnica del fenómeno, fortaleciendo la tutela judicial efectiva, la prevención general y la coherencia en la motivación de las sentencias.

Propuesta metodológica para aplicar la matriz en expedientes reales

Esta propuesta convierte el marco conceptual y los criterios operativos en un procedimiento reproducible para fiscales, jueces y peritos cuando la IA fue el medio del delito. Se apoya en tres premisas del manuscrito: (i) autor humano e IA como instrumento; (ii) dosificación guiada por el daño real y la gravedad del resultado; y (iii) un esquema orientador —no un cuadro punitivo cerrado— que mejore la proporcionalidad y el efecto preventivo.

Orientadores del proceso

Se propone que el expediente penal integre y documente los siguientes elementos:

a) **Evidencia técnica del uso de IA.** Archivo original, hash, metadatos y logs, con preservación y cadena de custodia.
b) **Evidencia de alcance.** Datos de visualizaciones, reenvíos y posicionamiento en buscadores; cobertura en medios/redes y replicabilidad.
c) **Evidencia de impacto.** Peritajes psicológicos y económicos; constancias laborales y familiares que acrediten consecuencias.

d) **Medidas de mitigación.** Retiros (*takedowns*), rectificaciones y desindexación, con constancias de gestión y eficacia.

e) **Dimensiones y gravedad del daño.** Valoración motivada (baja/media/alta/muy alta) en: moral/reputacional, integridad física/psicológica, económico-patrimonial, profesional, familiar/social y privacidad/intimidad.

f) **Agravantes (uso de IA).** Verosimilitud sintética alta (p. ej., *deepfakes*); automatización y escala de difusión; opacidad/ocultación (VPN/TOR, identidades ficticias, borrado de metadatos); persistencia del contenido; vulnerabilidad de la víctima; campaña reiterada; impacto laboral probado; lucro; afectación a terceros; riesgo a la integridad física; asociación para delinquir.

g) **Atenuantes (condicionados a verificación).** Colaboración con la investigación (especialmente en asociación para delinquir o redes); culpa leve y ausencia de ganancia; hecho aislado sin automatización; discapacidad o capacidad disminuida debidamente acreditada.

> **Regla de motivación:** cada elemento debe constar expresamente en la causa y su valoración ha de impactar, de forma razonada, la ubicación de la pena dentro de la banda aplicable.

Procedimiento (8 fases, con entregables)

a) Admisión y cadena de custodia *(ver Anexos 1 y 7)*

- Apertura de oficio o a instancia de parte.
- Ficha de caso–IA: registro de fuentes y preservación de evidencias digitales.
- Entregable: Acta de preservación (hashes, rutas, custodios).

b) Verificación estructural *(ver Anexo 2)*

- Constatar la existencia de un autor humano e IA como medio.
- Determinar dolo o culpa.
- Entregable: Informe de verificación (hechos, herramienta de IA, rol causal).

c) Pericia técnica en IA *(ver Anexo 6)*

- Validación de síntesis (*deepfake/voice/text*), trazabilidad y alcance (bots/redes).
- Entregable: Dictamen técnico (método, hallazgos, límites).

d) Valoración del daño *(ver Anexo 3)*

- Oficios a entes policiales de investigación con capacidad y entrenamiento en IA.
- Peritajes psicológico, económico y social; recopilación de constancias laborales/familiares y de privacidad afectada; testigos; elementos electrónicos, publicitarios y digitales.
- Entregable: Tabla de daño.

e) Determinación de la gravedad *(ver Anexo 5)*

- Clasificar el resultado, con especial atención a la irreversibilidad (p. ej., pérdida reputacional/profesional duradera).
- Entregable: Acta de gravedad.

f) Agravantes y atenuantes (AA) *(ver Anexo 4)*

- Aplicar el checklist específico de IA.
- Entregable: Matriz AA.

g) Integración y dosificación *(ver Anexo 8)*

- Combinar Daño × Gravedad + AA para proponer el tramo dentro del rango legal del tipo.
- Entregable: Informe de dosificación motivada (reglas, pruebas y fundamentos).

h) Control de calidad y revisión por pares *(ver Anexo 9)*

- Revisión por segundo fiscal/perito (doble firma) para verificar consistencia y proporcionalidad.
- Entregable: Acta de conformidad.

Instrumentos estandarizados

- **Ficha Caso–IA (Anexo 1):** datos de identificación, descripción de hechos, herramientas empleadas, fuentes y artefactos recolectados, cadena de custodia (hashes, metadatos, custodios y transferencias).
- **Checklist de Verificación Estructural (Anexo 2):** confirmación de autor humano, IA como medio, y determinación de dolo/culpa.
- **Hoja de Matriz Daño × Gravedad (Anexo 3):** valoración en seis dimensiones con categoría de resultado (baja/media/alta/muy alta) y motivación probatoria.
- **Matriz de Agravantes/Atenuantes – AA (Anexo 4): 10 agravantes / 5 atenuantes, con criterios de activación y evidencia mínima exigida.**
- **Plantilla de Motivación de Dosificación (Anexo 8):** integración de fundamentos fácticos, técnicos y jurídicos y selección del tramo dentro del rango legal del tipo.

> Todos los formatos deben adjuntarse al expediente, con firmas y fechas del responsable técnico y del fiscal/juez que corresponda.

Confiabilidad y validez

- **Capacitación y manual operativo.** Programa de formación para fiscales y peritos, con manual ilustrado que incluya casos tipo de difamación e injuria, procedimientos de preservación de evidencia digital y ejemplos de motivación de decisiones.
- **Doble calificación independiente.** Aplicación ciega y por duplicado de la Matriz Daño × Gravedad y de la Matriz AA; cálculo del coeficiente de acuerdo interevaluador y resolución de discrepancias mediante revisión técnica conjunta.
- **Plan Piloto con 20 expedientes.** Ensayo controlado para ajustar umbrales, depurar redacción de ítems, verificar consistencia y estimar tiempos de aplicación y carga probatoria.
- **Trazabilidad probatoria.** Cada inferencia de daño o gravedad se ancla a evidencia identificada y foliada en el expediente (hashes, metadatos, logs, analíticas, peritajes), con referencias a anexos y cadena de custodia.

Motivación y proporcionalidad (modelo de redacción)

La motivación debe articular, de forma clara y verificable, cinco bloques. A continuación se ofrece un modelo operativo listo para insertar en resoluciones o informes.

a) **Hecho y medio IA**

Fórmula sugerida:

"Se tiene por acreditado que [A], en fecha [dd/mm/aaaa], ejecutó la conducta consistente en [describir hechos], utilizando inteligencia artificial como medio instrumental, específicamente [herramienta/función: p. ej., síntesis de voz/deepfake/generación de texto], lo que potenció el alcance, la verosimilitud y la velocidad de difusión del contenido."

b) **Daño y gravedad (Matriz Daño × Gravedad**

Fórmula sugerida:

"De los peritajes y constancias obrantes se establece un daño [moral/reputacional, físico/psicológico, económico-patrimonial, profesional, familiar/social, privacidad] con gravedad [baja/media/alta/muy alta], atendida la [persistencia/irreversibilidad/replicabilidad] del contenido y su alcance [número de visualizaciones/reenvíos/cobertura].

En la Matriz Daño × Gravedad el caso se ubica en la categoría [indicar]."

c) **Agravantes / Atenuantes (AA)**

Fórmula sugerida:

"Concurren las agravantes [verosimilitud sintética alta; automatización/escala; opacidad; persistencia; vulnerabilidad; reiteración; impacto laboral; lucro; afectación a terceros; riesgo a la integridad; asociación], debidamente probadas en [folios/anexos].

Se aprecia la atenuante [retractación eficaz/retirada y desindexación diligentes/cooperación técnica/reparación integral/rol marginal/etc.], cuya eficacia se acredita en [folios/anexos]."

d) **Selección de tramo dentro del rango legal del tipo**

Fórmula sugerida:

"Partiendo de la pena base orientativa del tipo [difamación/injuria], la concurrencia de [agravantes] y el perfil de daño [alto/muy alto], con [irreversibilidad/persistencia], desplaza la respuesta al tramo [medio/alto] del rango legal. En consecuencia, se propone/impone [pena concreta], junto con medidas complementarias de rectificación equivalente, reparación y no repetición." (Anexo 8: Informe de dosificación motivada).

e) **Finalidad preventiva (proporcionalidad y visibilidad)**

Fórmula sugerida:

"La proporcionalidad de la pena —visiblemente ajustada al daño real y a la gravedad del resultado cuando la IA es el medio comisivo— fortalece la tutela judicial efectiva y eleva el efecto disuasorio general. En la medida en que el sistema haga explícita esta correlación y la publicidad sancionatoria sea idónea, es razonable esperar una reducción de la incidencia de estas conductas por mayor visibilidad y previsibilidad de la respuesta penal."

Notas de estilo para la motivación

- Siempre anclar cada inferencia a evidencia identificada (folio, anexo, hash, peritaje).
- Evitar fórmulas genéricas: motivar dimensión por dimensión de la matriz.
- Distinguir con precisión IA como medio (no autor) y sujeto activo humano.
- Explicar por qué las medidas restaurativas (retiro, desindexación, rectificación) fueron o no equivalentes al alcance del daño.
- Incorporar, cuando proceda, medidas complementarias: rectificación equivalente, programas de gestión reputacional, atención psicológica y monitoreo de no repetición.

Ejemplo sintético (ilustrativo)

"Consta acreditado que A difundió un video deepfake que atribuía a B conductas deshonrosas (medio IA: síntesis de imagen y voz). La difusión alcanzó [X] visualizaciones y [Y] reenvíos en [plataformas], con replicación multicanal e índices de permanencia elevados (Anexos 1, 3 y 6).

El daño moral/reputacional y profesional se califica alto, con irreversibilidad media-alta por persistencia en buscadores (Anexo 5). Concurren agravantes de verosimilitud sintética, automatización y reiteración; la retractación ofrecida careció de equivalencia y alcance (Anexo 4). Se ubica la pena en el tramo alto del tipo de difamación, con [pena] y medidas de rectificación equivalente, reparación integral y no repetición (Anexo 8). La decisión busca reforzar el efecto preventivo mediante proporcionalidad motivada."

Consideraciones éticas y de debido proceso

Este artículo se limita a proponer premisas y herramientas de dosificación. No crea sanciones nuevas ni pretende constituir las reformas legales que, en su caso, correspondan.

Marcos internacionales

- **Unión Europea —** AI Act. En vigor desde el 1 de agosto de 2024; aplicación plena el 2 de agosto de 2026, con fases intermedias: prohibiciones y alfabetización en IA desde 2 feb. 2025; obligaciones para GPAI y gobernanza desde 2 ago. 2025; prórroga hasta 2 ago. 2027 para ciertos sistemas de alto riesgo. Este calendario crea incentivos de trazabilidad técnica y gobernanza que inciden en la prueba digital y la persecución penal. (Digital Strategy EU).
- **Consejo de Europa —** Convenio Marco sobre IA (CETS 225). Primer tratado jurídicamente vinculante en la materia; abierto a firma el 5 de septiembre de 2024. Exige que todo el ciclo de vida de la IA sea compatible con derechos humanos, democracia y Estado de derecho; varios países (incluidos EE. UU., UE y Reino Unido) han firmado el instrumento. (Portal).
- **OCDE —** Principios de IA. Estándar intergubernamental (2019, actualizado en mayo de 2024) que aporta principios de proporcionalidad, responsabilidad y transparencia, y recomendaciones de política para una IA confiable; la actualización busca interoperabilidad global ante los modelos generativos. (oecd.ai).
- **UNESCO —** Recomendación sobre la Ética de la IA. Estándar global (2021; última actualización sept. 2024) aplicable a 194 Estados Miembros; subraya derechos humanos, equidad y supervisión humana. (UNESCO).
- **Nota transversal.** Aunque no son normas penales “puras”, estos marcos fijan parámetros transversales (rendición de cuentas, trazabilidad, gestión de riesgos) útiles para valorar agravantes vinculadas a verosimilitud sintética, automatización y opacidad técnica.

- **Convenio de Budapest sobre Ciberdelincuencia** y su **Segundo Protocolo Adicional**. Establecen reglas para mejorar la cooperación y el acceso transfronterizo a evidencia electrónica; crítico cuando el contenido ilícito generado con IA se aloja fuera de la jurisdicción. (Portal).
- **Estados Unidos — NIST AI Risk Management Framework (AI RMF 1.0).** Guía voluntaria (enero 2023) para gestionar riesgos de IA; en 2024 NIST publicó un perfil para IA generativa, alineado con la Orden Ejecutiva 14110 (30 oct. 2023) sobre desarrollo seguro, confiable y responsable, que instruye a agencias y estándares técnicos. (nvlpubs.nist.gov).
- **Reino Unido y socios — Bletchley Declaration (nov. 2023).** Compromiso de 28 países para cooperar en evaluación de riesgos y pruebas de seguridad de modelos avanzados; Reino Unido y EE. UU. anunciaron colaboración bilateral para ensayar modelos de frontera. (GOV.UK).
- **Unión Europea (defensa y seguridad).** Documentos técnicos de la UE destacan que el AI Act adopta un enfoque por riesgo y definen ámbitos excluidos (p. ej., usos militares), útil para trazar fronteras regulatorias al analizar expedientes. (eda.europa.eu).
- **Brasil.** El Senado aprobó el PL 2338/2023 (dic. 2024), que establece un marco de gobernanza, transparencia y sanciones (multas hasta BRL 50 millones o 2 % de ingresos), aún pendiente de tramitación final. (mattosfilho.com.br).
- **Chile.** Cuenta con Política Nacional de IA 2021–2030 y, en 2024, presentó actualización y proyecto de ley; el eje ético-regulatorio se alinea con la Recomendación de UNESCO. (oecd.ai).

- **México.** Fuentes internacionales reportan un panorama mixto: estrategia nacional en 2018, pero UNESCO señala que actualmente no hay un plan/estrategia vigente; existen propuestas 2024–2030 e iniciativas legislativas en curso. *Implicación: oportunidad de armonizar con AI Act/CoE/OCDE/UNESCO y con instrumentos de evidencia electrónica*. (s899a9742c3d83292.jimcontent.com).

- **Aceleración delictiva con IA.** Organismos europeos y foros internacionales subrayan que la IA potencia, acelera y sofistica la comisión de ilícitos y dificulta su persecución (Interpol), lo que exige nuevos criterios de trazabilidad y evaluación de riesgos, coherentes con el AI Act y los Principios OCDE. (Digital Strategy EU).

- **Convergencia en *risk-based governance*.** Los instrumentos citados privilegian gestión de riesgos, responsabilidad y debida diligencia (p. ej., evaluación de impacto, pruebas de seguridad, logs y metadatos), reforzando en nuestra metodología los calibradores probatorios (huellas digitales, alcance de difusión, persistencia en línea). (nvlpubs.nist.gov).

- **Armonización internacional en curso.** El AI Act y el Convenio del CoE avanzan, pero no existe un consenso global consolidado; este vacío crea un espacio de oportunidad regulatoria para países que se adelanten. (White & Case).

- **Situación venezolana.** Se identifica un desfase entre el daño real causado con IA en difamación/injuria y los rangos penales vigentes, lo que debilita la prevención general.

Discusión y aportes

Valor agregado de la investigación

Aporte central. Integra un marco conceptual claro (autor humano; IA como medio) con una metodología operativa de dosificación basada en Daño × Gravedad, más agravantes/atenuantes específicos del uso de IA. Con ello, traduce debates abstractos en herramientas aplicables por fiscalías y jueces, sin pretender fijar un cuadro punitivo cerrado.

Novedad práctica. La Matriz Daño × Gravedad obliga a visualizar el impacto real y no solo el tipo penal formal, lo que podría justificar la elaboración de una ley especial con sanciones más altas cuando la IA sea el medio delictivo, en atención a la persistencia, alcance e irreversibilidad del daño.

Valor metodológico y replicabilidad

- **Estandarización.** Fichas, *checklists* y actas crean un lenguaje común entre peritos, fiscales y jueces.
- **Motivación robusta.** Véase Anexo 8.
- **Revisión por pares.** La doble firma agrega consistencia y habilita auditoría metodológica (véase Anexo 9).

Conexión con estándares internacionales

El enfoque dialoga con el AI Act (gestión de riesgos, trazabilidad), el Convenio Marco del Consejo de Europa, OCDE y UNESCO (responsabilidad, transparencia, debida diligencia). Operativamente, esto respalda tratar como agravantes la verosimilitud sintética, la automatización/escala, la opacidad técnica y la persistencia. La referencia al Convenio de Budapest +

Segundo Protocolo refuerza diligencias de evidencia electrónica transfronteriza.

Implicaciones para Venezuela

La metodología corrige la subvaloración punitiva de la difamación/injuria mediada por IA, fortalece la prevención general e incorpora criterios operativos a los programas de formación del Ministerio Público y del Poder Judicial, con guías de uso y casuística (véanse Casos y Anexos).

Efectos esperados en la práctica

Las variables propuestas mejoran la calidad probatoria mediante el uso sistemático de dictámenes técnicos de IA, actas de preservación y cadenas de custodia (Anexos 6–7), reducen la discrecionalidad con matrices y plantillas de motivación (Anexos 3–4 y 8) y aumentan la coherencia intercasos (Anexo 9). En conjunto, ello previsiblemente desplaza la pena hacia tramos superiores cuando el daño sea alto y refuerza el efecto preventivo del sistema.

Recomendaciones operativas (implementación inmediata)

a) Protocolizar el uso de la Matriz Daño × Gravedad y del checklist de IA en todo expediente con indicios tecnológicos (véanse Anexos 3–4).
b) Exigir pericias técnicas mínimas —autenticidad, alcance y automatización— antes de dosificar (véase Anexo 6).
c) Aplicar revisión por pares en decisiones de tramo alto o cuando el daño se califique como “extremo” (véase Anexo 9).

Indicadores para evaluar impacto (seguimiento)

- % de expedientes con uso completo de Anexos 3–4–5–8.
- Tiempo de tramitación pericial en IA (línea base vs. post-implementación).
- Tasa de confirmación de la dosificación en alzada.
- Medidas de reparación: ordenadas vs. efectivamente cumplidas.
- Reducción de reincidencia en tipologías con IA como medio.

Limitaciones y líneas futuras

Limitaciones

El presente artículo es orientador y se circunscribe a Venezuela, priorizando difamación e injuria con IA como medio. Aunque el marco es trasladable, su validez externa hacia otros tipos penales requiere pruebas adicionales.

Su implementación —sea vía reforma de leyes vigentes o creación de una ley especial— demanda capacitación, laboratorios forenses, acceso a software y peritajes especializados, además de protocolos de cadena de custodia digital. Sin estos recursos, la aplicación efectiva de cualquier actualización normativa será inviable.

Debe considerarse la cooperación transfronteriza, pues sistemas y servidores pueden estar fuera del país; la falta de canales ágiles de cooperación internacional limita la obtención de evidencia clave. Asimismo, existen asimetrías técnicas (dependencia de plataformas para *logs* y metadatos), posibles cuellos de botella periciales y el riesgo de respuestas punitivas desproporcionadas si no se respeta el debido proceso y la motivación técnica caso por caso.

Líneas futuras

1. **Planes pilotos controlados.** Implementar pruebas con 20–50 expedientes para ajustar umbrales, tiempos, redacción de ítems y cargas probatorias de la matriz *Daño* × *Gravedad* y la Matriz AA.
2. **Protocolos nacionales de evidencia digital.** Estandarizar preservación, hashing, metadatos, *logs*, trazabilidad de difusión y validación de deepfakes/clonación de voz; checklist único para MP, PJ y cuerpos de investigación.
3. **Capacitación y certificación.** Diseñar currículos para fiscales, jueces y peritos (cadena de custodia, herramientas IA, límites técnicos, motivación de decisiones), con certificación periódica.
4. **Convenios y puntos de contacto 24/7.** Fortalecer cooperación con plataformas y canales internacionales para acceso a evidencia electrónica y medidas urgentes de retiro/desindexación.
5. **Jurisprudencia y casuística.** Crear un repositorio de decisiones y casos tipo (anónimo), con boletines de buenas prácticas y criterios de proporcionalidad.
6. **Indicadores y tablero de control.** Medir uso de Anexos 3–4–5–8, tiempos periciales, confirmación en alzada, cumplimiento de medidas restaurativas y reincidencia.
7. **Armonización normativa.** Explorar reformas para incorporar el "daño causado" como elemento a considerar en la dosificación cuando medie IA, y agravantes vinculadas a verosimilitud sintética, automatización y persistencia.
8. **Protección integral de víctimas.** Protocolizar medidas cautelares ágiles (retiro, desindexación, rectificación equivalente, no repetición), atención psicológica y apoyo de gestión reputacional.

9. **Ética y derechos fundamentales.** Incorporar salvaguardas de libertad de expresión, privacidad y no discriminación, con ponderación reforzada para figuras públicas e interés público.

10. **Infraestructura forense.** Invertir en laboratorios, software validado y bancos de referencia (datasets en español/latam) para pruebas de detección y atribución.

11. **Gobernanza de riesgos institucional.** Adoptar marcos de gestión de riesgos para IA en las instituciones de justicia (políticas internas, *compliance*, auditorías técnicas).

12. **Investigación empírica.** Evaluar el impacto preventivo de la dosificación proporcional con series de tiempo y análisis intercasos, ajustando la metodología a evidencia.

Este itinerario permite operativizar la propuesta, reducir brechas técnicas y normativas, y avanzar hacia una respuesta penal proporcional y efectiva frente a la difamación e injuria cometidas con IA como medio.

Conclusiones y recomendaciones

Conclusiones

El propósito de este trabajo es sensibilizar sobre la necesidad de actualizar las normas que sancionan los delitos de difamación e injuria cuando se cometen utilizando la IA como medio, articulándolo con un debido proceso ajustado a la gravedad del daño causado. La propuesta convoca a fiscales, jueces y peritos a dosificar con mayor proporcionalidad y mejor motivación.

Nuestro aporte se funda en la incorporación expresa de tres elementos esenciales en la construcción del caso y su dosificación: (i) la dignidad como bien jurídico tutelado y eje unificador de los demás derechos de la persona; (ii) la IA como medio comisivo en cualquiera de sus manifestaciones digitales; y (iii) el daño causado, que exige valorar dimensiones morales, económicas, profesionales, familiares, de intimidad y privacidad, así como la integridad y la irreversibilidad digital.

Sostenemos, además, la conveniencia de adaptar el proceso para que el Ministerio Público pueda obrar de oficio en defensa de la dignidad protegida por la Constitución, fortaleciendo la prevención general y corrigiendo la subvaloración punitiva que hoy suele observarse.

Dado el carácter dinámico del fenómeno —con parámetros técnicos que se expanden a gran velocidad—, el sistema de justicia requiere formación continua en IA: trazabilidad probatoria, pericia técnica, cadena de custodia digital y dosificación. La ciencia jurídica debe actualizarse de forma permanente para asegurar una aplicación de justicia efectiva.

En el plano internacional, es imprescindible atender a la gobernanza basada en riesgos (UE/OCDE/UNESCO) y a la cooperación para evidencia electrónica (Convenio de Budapest), tanto por la novedad regulatoria como por la localización transfronteriza de servicios, servidores y nubes. Ello ofrece un puente práctico entre el expediente local y los marcos comparados, y habilita acuerdos de colaboración para la obtención de prueba y preservación de la cadena de custodia.

Finalmente, aunque el estudio se centra en difamación e injuria, sus lineamientos son trasladables a otros tipos —calumnia, estafa, delitos de odio, entre otros— cuando la IA sea el medio.

Recomendaciones

1. Implementación inmediata de la matriz Daño × Gravedad, el checklist de IA y las plantillas de motivación en expedientes con indicios tecnológicos.

2. Capacitación sistemática de operadores (policiales, fiscales, judiciales y peritos) y dotación de laboratorios y software forense validado.

3. Protocolos nacionales de preservación de evidencia digital, hashing, metadatos y verificación de deepfakes/clonación de voz.

4. Cooperación internacional: puntos de contacto y acuerdos para acceso transfronterizo a evidencia y medidas urgentes de retiro/desindexación.

5. Revisión por pares (doble firma) en decisiones de tramo alto o con daño extremo, y tablero de indicadores para seguimiento y mejora continua.

Referencias

AP News. (2025, 18 de marzo). AI is turbocharging organized crime, EU police agency warns. *AP News*. https://apnews.com/article/europol-ai-organized-crime-641b0b0a0b8d4f2f-9b2a1db9e0d0c1a2

Autor(es). (2025). *Dosificación penal de los delitos cometidos mediante IA* [Manuscrito inédito, PDF del autor]. O! Ediciones.

Brasil. Senado Federal. (2024, 10 de diciembre). *PL 2338/2023* (aprobación en Senado; en trámite en la Cámara de Diputados). https://www25.senado.leg.br/web/actividad…/materia/163566

Cámara de Diputadas y Diputados de Chile. (2024, 7 de mayo–). *Proyecto de ley que regula los sistemas de IA* (Boletín 16821-19) [Tramitación]. https://www.camara.cl/legislacion/ProyectosDeLey/tramitacion.aspx?prmBOLETIN=16821&prmID=17429

Chile. Ministerio de Ciencia, Tecnología, Conocimiento e Innovación. (2021). *Política Nacional de Inteligencia Artificial*. https://minciencia.gob.cl/areas/inteligencia-artificial/politica-nacional-de-inteligencia-artificial/

Chile. Ministerio de Ciencia, Tecnología, Conocimiento e Innovación. (2024). *Política Nacional de Inteligencia Artificial (Actualización 2024)*. https://cens.cl/wp-content/uploads/2024/05/Politica-Nacional-de-IA-Actualizada-2-05.pdf

Comisión Europea. (2024, 1 de agosto). *AI Act enters into force*. https://commission.europa.eu/news-and-media/news/ai-act-enters-force-2024-08-01_en

Comisión Europea – DG CONNECT. (2025). *Regulatory framework for AI (AI Act): Application timeline*. https://digital-strategy.ec.europa.eu/en/policies/regulatory-framework-ai

Consejo de Europa. (2022). *Second Additional Protocol to the Convention on Cybercrime on enhanced co-operation and disclosure of electronic evidence*. https://www.coe.int/en/web/cybercrime/second-additional-protocol

Consejo de Europa. (2024). *Framework Convention on Artificial Intelligence and human rights, democracy and the rule of law (CETS 225)*. https://www.coe.int/en/web/artificial-intelligence/the-framework-convention-on-artificial-intelligence

Covington. (2024, 21 de noviembre). *International CoE Convention on AI—What to expect after its signature?* https://www.cov.com/-/media/files/corporate/publications/2024/11/international-coe-convention-on-aiwhat-to-expect-after-its-signature.pdf

Eurojust. (2024, 18 de octubre). *Article 9 of the Second Additional Protocol… (expedited disclosure).* https://www.eurojust.europa.eu/publication/article-9-second-additional-protocol-convention-cybercrime-expedited-disclosure-stored

Europol. (2024). *AI and policing: The benefits and challenges of artificial intelligence for law enforcement*. https://www.europol.europa.eu/publication-events/main-reports/ai-and-policing

Europol Innovation Lab. (2023). *Facing reality? Law enforcement and the challenge of deepfakes*. https://www.europol.europa.eu/publication-events/main-reports/facing-reality-law-enforcement-and-challenge-of-deepfakes

Gobierno del Reino Unido. (2023, 1–2 de noviembre). *The Bletchley Declaration by countries attending the AI Safety Summit*. https://www.gov.uk/government/publications/ai-safety-summit-2023-the-bletchley-declaration/the-bletchley-declaration-by-countries-attending-the-ai-safety-summit-1-2-november-2023

NIST. (2023). *Artificial Intelligence Risk Management Framework (AI RMF 1.0)* (NIST AI 100-1). National Institute of Standards and Technology. https://nvlpubs.nist.gov/nistpubs/ai/NIST.AI.100-1.pdf

NIST. (2024). *AI RMF Generative AI Profile* (NIST AI 600-1). National Institute of Standards and Technology. https://nvlpubs.nist.gov/nistpubs/ai/NIST.AI.600-1.pdf

OCDE. (2024, 3 de mayo). *AI Principles* (actualización 2024). https://oecd.ai/en/ai-principles

Parlamento Europeo (EPRS). (2025, 10 de junio). *AI Act implementation timeline (At a Glance)* (EPRS_ATA(2025)772906). https://www.europarl.europa.eu/thinktank/en/document/EPRS_ATA(2025)772906

Estados Unidos. Federal Register. (2023, 1 de noviembre). *Safe, secure, and trustworthy development and use of artificial intelligence* [Executive Order 14110]. https://www.federalregister.gov/documents/2023/11/01/2023-24283/safe-secure-and-trustworthy-development-and-use-of-artificial-intelligence

Estados Unidos. La Casa Blanca. (2023, 30 de octubre). *Executive Order 14110: Safe, Secure, and Trustworthy Development and Use of Artificial Intelligence*. https://bidenwhitehouse.archives.gov/briefing-room/presidential-actions/2023/10/30/executive-order-on-the-safe-secure-and-trustworthy-development-and-use-of-artificial-intelligence/

UNESCO. (2024, 4 de julio). *Artificial Intelligence Readiness Assessment – Mexico (RAM)*. https://www.unesco.org/ethics-ai/en/articles/unesco-presents-artificial-intelligence-readiness-assessment-mexico

UNESCO. (2024, 26 de septiembre). *Recommendation on the Ethics of Artificial Intelligence* (actualizada). https://www.unesco.org/en/artificial-intelligence/recommendation-ethics

White & Case. (2024, 16 de julio). *Long-awaited EU AI Act becomes law after publication in the EU's Official Journal*. https://www.whitecase.com/insight-alert/long-awaited-eu-ai-act-becomes-law-after-publication-eus-official-journal

Anexos

Anexo 1:

Ficha Caso IA – Apertura y cadena de custotia

1. Datos generales

Nº de expediente: _____ | Fiscalía/Juzgado: _____

Víctima(s): _____

Imputado(s): _____

__

Tipo penal presunto: □ Difamación □ Injuria □ Estafa □ Otro:

_____

Fecha y lugar del hecho: _____

2. Descripción breve del hecho

Relato sintético (máx. 8 líneas):

3. Rol de la IA (medio/instrumento)

Herramienta o técnica: □ Deepfake video □ Clonación de voz □ Texto IA □ Generador de imágenes □ Automatización/bots □ Otro:

_____

Función en la ejecución: _____

4. **Evidencia digital y cadena de custodia**

Evidencia	Descripción	Origen/URL	Hash	Custodio	Fecha/Hora
1					
2					

5. Medidas tempranas de preservación

□ Descarga forense □ Capturas con metadatos □ Notificación a plataforma

□ Solicitud de takedown □ Otra: _____

Anexo 2:

Checklist verificación estructural

A. Requisitos de entrada

□ Sujeto activo humano identificado (al menos indiciariamente).

□ IA utilizada como medio/instrumento del hecho (descrito y vinculado).

□ Nexo causal entre uso de IA y resultado (alcance/efecto).

□ Calificación del elemento subjetivo: □ Dolo □ Culpa (especificar).

Observaciones: _____

B. Pertinencia probatoria mínima

□ Evidencia técnica del uso de IA (pericia/indicio técnico).

□ Evidencia de alcance (métricas, SEO, reenvíos).

□ Evidencia de impacto (pericias psicológicas/económicas, constancias).

□ Registro de medidas de mitigación (takedown, rectificaciones).

Resultado:

□ Procede aplicar Matriz Daño×Gravedad □ Subsanar (especificar)

Anexo 3:

Hoja matriz Daño × Gravedad

Dimensión del daño	Evidencia/Referencia
Moral/Reputacional	
Integridad física (víctima/terceros)	
Económico/Patrimonial	
Profesional/Laboral	
Familiar/Social	
Privacidad/Datos sensibles	

Anexo 4:

Matriz Agravantes/Atenuantes

Agravantes (sume 1 por cada casilla marcada):

□ Verosimilitud sintética alta (deepfake/voice/text persuasivo)

□ Automatización/escala (bots, schedulers, gran difusión)

□ Opacidad/ocultación (VPN/TOR, sockpuppets, borrado metadatos)

□ Persistencia/dificultad de remoción/desindexación

□ Vulnerabilidad de la víctima (menor/no figura pública/dependencia)

□ Reiteración/campaña coordinada

□ Impacto laboral probado (pérdida empleo/contratos)

□ Aprovechamiento lucrativo

□ Afectación a terceros (familia, clientes)

□ Riesgo a integridad física (acoso, amenazas)

Total agravantes (0–10): _____

Atenuantes (reste 1 por cada casilla marcada):

□ Retractación y retiro diligente del contenido

□ Reparación pronta e integral (limpieza reputacional, terapia, costos legales)

□ Colaboración con la investigación (herramientas/cuentas)

□ Culpa leve/sin ganancia/ausencia de intención

□ Hecho aislado sin automatización/plan

□ Persona con incapacidad o limitación o condición especial

Total atenuantes (0–5): _____

Documente la prueba que respalda cada marca.

Anexo 5:

Acta de gravedad del resultado

Caso: _____ Fecha: _____

Evaluadores: _____

1. Resumen del daño probado (3–5 líneas): _____
2. Clasificación final de Gravedad (marque): □ 0 □ 1 □ 2 □ 3
3. Justificación (vincule con evidencias): _____

Firma(s): _____

Anexo 6:

Dictamen técnico de IA (Plantilla pericial)

Encargo: autenticidad/manipulación sintética, trazabilidad, alcance.

Perito: _____ Acreditación: _____

Metodología (herramientas, versiones, límites):

_____

Hallazgos técnicos

Indicios de síntesis/manipulación:

_____

Metadatos/Hashes: _____

Automatización/bots/redes: _____

Alcance y persistencia: _____

Limitaciones (sesgos, falsos positivos, cobertura): _____

Conclusión técnica (en lenguaje llano, 5–8 líneas):

Anexo 7:

Acta de preservación de evidencia digital

Fuentes preservadas: (archivos, URLs, IDs de publicaciones)

Ítem	Fuente/URL	Fecha/hora captura	Hash	Custodio	Observaciones
1					
2					

Procedimientos aplicados: □ Captura forense □ Wget/HTTrack □ API/Export

□ Otro: _____

Observaciones de integridad: _____

Firmas: _____ / _____

Anexo 8:

Informe de dosificación motivada

(Modelo deredacción)

1. Hechos y medio IA

Describir la conducta y cómo la IA operó como instrumento del delito.

2. Matriz Daño × Gravedad

Transcribir el promedio de daño y la categoría de gravedad, con referencia a pericias/constancias.

3. Agravantes/Atenuantes

Listar los ítems marcados, su evidencia y el balance neto.

4. Selección de tramo dentro del rango legal del tipo

Tipo penal: _____ Rango legal: _____

Tramo seleccionado: □ Bajo □ Medio □ Alto

Fundamento (por qué ese tramo es proporcional y adecuado).

5. Finalidad preventiva y restaurativa

Breve explicación de cómo la dosificación propuesta contribuye a disuasión y reparación.

Anexo 9
Acta de conformidad
(Revisión par)

Revisor/a: _____ Fecha: _____

Cotejo realizado:

□ Consistencia entre hechos, prueba técnica y matriz

□ Motivación suficiente de agravantes/atenuantes

□ Proporcionalidad del tramo seleccionado

□ Lenguaje claro y respetuoso de garantías

Observaciones y ajustes sugeridos:

_____

Resultado: □ Conforme □ Conforme con observaciones □ Requiere subsanar

Firma: _____

Anexo 10:

Plantilla de “Checks” rápidos para audiencias

ü ¿Autor humano + IA como medio descritos con claridad? □

ü ¿Daño probado o una crítica con evidencia sólida? □

ü ¿Gravedad justificada con hechos y pericias? □

ü ¿Agravantes/atenuantes sustentados en prueba específica? □