Deepfakes, inteligencia artificial generativa y derecho a la integridad digital: análisis crítico del proyecto de ley chileno sobre imitaciones no consentidas (Boletín N°17.795-19)

Deepfakes, Generative Artificial Intelligence, and the Right to Digital Integrity: A Critical Analysis of the Chilean Bill on Non-Consensual Imitations (Bill No. 17,795-19)

Felipe Diez Ringele*

Universitat de Girona, Girona, España - Universidad del Desarrollo, Santiago, Chile

fdiezr@udd.cl | http://orcid.org/0009-0007-1891-5574

◆

Recibido: 15/12/2025 | Aceptado: 10/04/2026 | Publicado: 08/05/2026

Resumen. La inteligencia artificial generativa ha posibilitado la creación de *deepfakes*, imitaciones digitales hiperrealistas de rostros, voces y cuerpos que pueden difundirse sin consentimiento y generar riesgos significativos para la identidad personal, la privacidad y la seguridad. Frente a este fenómeno, Chile presentó en 2025 un proyecto de ley destinado a reconocer un derecho exclusivo sobre la integridad digital y a regular el uso no autorizado de estas tecnologías. Este artículo analiza (i) los fundamentos técnicos y jurídicos de los *deepfakes*, (ii) las definiciones y alcances del proyecto de ley, y (iii) el régimen de responsabilidades y sanciones propuesto, evaluando sus implicancias regulatorias y su adecuación al contexto normativo contemporáneo.

Palabras clave. *deepfakes*; inteligencia artificial generativa; identidad digital; integridad digital; regulación de IA.

Abstract. Generative artificial intelligence has enabled the creation of deepfakes, hyper-realistic digital imitations of faces, voices, and bodies that can be disseminated without consent, posing significant risks to personal identity, privacy, and security. In response, Chile introduced in 2025 a bill aimed at recognizing an exclusive right to digital integrity and

* Abogado. Candidato a Doctor en Derecho por la Universitat de Girona (España) y la Universidad del Desarrollo (Chile). Máster en Derecho.

regulating the unauthorized use of these technologies. This article examines the technical and legal foundations of deepfakes, the definitions and scope of the proposed legislation, and its system of responsibilities and sanctions, assessing its regulatory implications and alignment with contemporary legal frameworks.

Keywords. deepfakes; generative artificial intelligence; digital identity; digital integrity; AI regulation.

Introducción

El término *deepfake* combina *deep* (profundo) y *fake* (falso) y designa contenidos generados mediante "*deep*" *learning* (aprendizaje profundo) que reproducen características biométricas de personas, generalmente sin su consentimiento. Este tipo de aprendizaje automático, basado en redes neuronales profundas, experimentó su despegue a partir de 2011, alcanzando rendimientos iguales o superiores al desempeño humano en tareas como visión por computador, traducción, diagnóstico médico y juegos complejos. Su desarrollo ha sido posible gracias al uso de *hardware* altamente especializado, grandes volúmenes de datos y avances algorítmicos, consolidando a la inteligencia artificial como un motor de transformación tecnológica que plantea desafíos regulatorios, sociales y jurídicos sin precedentes (Russell y Norvig, 2022, pp. 44-48; Patil et al., 2023, p. 1).

En Chile, el 21 de agosto de 2025 se ingresó el proyecto de ley destinado a proteger a las personas frente al uso no consentido de inteligencia artificial para imitar su imagen, cuerpo o voz (Boletín 17.795-19)[1]. La iniciativa responde al diagnóstico de que la IA generativa permite crear *deepfakes* hiperrealistas capaces de replicar identidades sin autorización, generando riesgos como fraude, pornografía no consentida, desinformación, extorsión y daños psicológicos (Comisión Europea, 2025, pp. 1-2). A diferencia de otros ordenamientos que han desarrollado regímenes específicos sobre identidad digital, el derecho chileno carece de una regulación expresa: el derecho a la imagen solo se reconoce de forma

[1] Un reciente artículo sostiene que, aun a falta de ley, existen respuestas parciales en el ordenamiento jurídico: en el plano constitucional, el recurso de protección permite medidas urgentes frente a la vulneración de derechos como la honra y la vida privada, pero no asegura una reparación integral; en materia penal, los tipos vigentes (injurias, arts. 161-C y 161-D del Código Penal) resultan insuficientes por exigir captación real o por la ambigüedad frente a contenidos sintéticos, lo que evidencia un déficit de tipicidad; en cuanto a las leyes de violencia de género, la Ley 21.675 permite subsumir estos hechos como violencia simbólica o digital, aunque sin tipificación expresa; y, finalmente, la Ley 21.719 sobre protección de datos ofrece una vía indemnizatoria por tratamiento ilícito de datos personales (imagen, biometría), pero con limitaciones prácticas como la necesidad de una sanción previa y la dificultad de identificar responsables (Jabbaz, 2025, pp. 35-73).

implícita[2] y la normativa vigente no aborda adecuadamente ni la generación autónoma de imitaciones realistas ni los deberes de las plataformas digitales frente a estos contenidos[3].

El proyecto parte de la premisa de que la imagen, el cuerpo y la voz constituyen elementos esenciales de la identidad personal que, dada su fácil reproducción mediante IA generativa, requieren una protección jurídica específica. La creación de réplicas hiperrealistas basadas en datos biométricos, sin intervención ni control del titular, supone riesgos relevantes para la vida privada, la honra y la integridad moral y patrimonial, que no son cubiertos de manera suficiente por los actuales marcos de protección de datos o de propiedad intelectual. Su fundamento jurídico se encuentra en las garantías constitucionales de vida privada y honra (artículo 19 n.º 4 y 5), así como en la jurisprudencia que ha reconocido la imagen corporal como un atributo jurídicamente protegido.

Sobre esta base, la iniciativa articula una idea matriz clara: reconocer a cada persona un derecho exclusivo a controlar la creación y difusión de imitaciones digitales realistas de su imagen, cuerpo y voz generadas mediante IA. Para ello, exige consentimiento previo, prohíbe y sanciona los usos no autorizados, impone deberes específicos a las plataformas para la prevención, detección y retiro de contenidos ilícitos, y contempla mecanismos de reparación para las víctimas, en coordinación con la normativa de protección de datos personales.

En este contexto, el objetivo de este trabajo es analizar: que es la inteligencia artificial generativa (1); las imitaciones no consentidas o *deepfakes* (2); los objetos y definiciones del proyecto de ley (3); el emergente derecho a la integridad digital (4); y el régimen de responsabilidades y sanciones previsto (5), finalizando con nuestras conclusiones (6).

1. Inteligencia artificial generativa (GenAI)

La comprensión de la GenAI exige situarla dentro del entramado conceptual más amplio de la inteligencia artificial, distinguiendo entre la noción general de IA, su concreción en sistemas de inteligencia artificial y, más recientemente, en modelos y sistemas de uso general. En efecto, la evolución normativa y tecnológica ha desplazado el foco desde definiciones abstractas hacia categorías funcionales basadas en el riesgo y en el modo de despliegue de estas tecnologías. En este contexto, la GenAI no constituye una categoría aislada, sino una manifestación específica —particularmente sofisticada— de los modelos de IA de uso general, cuya capacidad de producir contenidos de manera autónoma y multimodal plantea desafíos jurídicos diferenciados, especialmente en materia de desinformación,

[2] Sobre el derecho a la propia imagen: Fernández, 2015.

[3] Con todo, podría sostenerse —desde una perspectiva crítica— que el ordenamiento jurídico chileno sí dispone de herramientas indirectas suficientes, en la medida en que la protección de los datos personales tiene reconocimiento constitucional y desarrollo en la Ley 19.628, pudiendo entenderse que la imagen, voz o rasgos biométricos constituyen datos personales en cuanto permiten identificar a una persona natural. Esta línea argumental encuentra cierto paralelo en el derecho comparado, particularmente en España, donde la Constitución Española de 1978 reconoce expresamente el derecho a la propia imagen (artículo 18.1 CE).

manipulación y afectación de derechos fundamentales. Desde esta perspectiva, resulta necesario examinar, en primer lugar, el concepto general de inteligencia artificial, para luego avanzar hacia sus concreciones normativas y, finalmente, situar adecuadamente la GenAI dentro de este marco.

1.1. Inteligencia artificial

La inteligencia artificial (IA) constituye un concepto abierto, lo que plantea una dificultad definitoria. Suele abordarse desde dos perspectivas principales: una antropocéntrica, que la relaciona con las facultades y capacidades propias del ser humano, y otra racionalista, que la asocia con la capacidad de actuar racionalmente. Sin embargo, ninguna de estas aproximaciones resulta plenamente satisfactoria, pues la naturaleza humana integra múltiples dimensiones estrechamente vinculadas a la inteligencia, y la propia racionalidad es un concepto tan complejo y esquivo como la inteligencia misma. En otras palabras, se intenta definir algo oscuro recurriendo a algo aún más oscuro (Amunátegui Perelló, 2021, pp. 13-14).

Asimismo, se ha sostenido que el concepto de IA es borroso, polifacético y dinámico. Es borroso porque no existe una definición consensuada, ni siquiera técnica, sobre qué condiciones debe cumplir una aplicación para ser considerada IA, lo que refleja la dificultad de precisar el propio concepto de inteligencia y el error de reducirlo a la capacidad de una máquina para simular conducta humana. Es polifacético, pues abarca diversos campos —como el *machine learning*, el *machine reasoning* o el *deep learning*— sin acuerdo sobre cuáles integran realmente la noción de IA. Y, es dinámico, porque su alcance varía con el avance tecnológico: lo que en su momento se consideró IA —como el filtrado de *spam*, la traducción automática o los motores de búsqueda— hoy se percibe como tecnología común (Bertolini, 2025, p. 37).

En lo que respecta a la Comisión Europea, esta ha señalado que el término "inteligencia artificial" (IA) se aplica a los sistemas que exhiben un comportamiento inteligente, en cuanto son capaces de analizar su entorno y actuar —con un cierto grado de autonomía— para alcanzar objetivos específicos. Asimismo, agrega que los sistemas de IA pueden consistir únicamente en un programa informático (por ejemplo, asistentes de voz, *software* de análisis de imágenes, motores de búsqueda o sistemas de reconocimiento facial y de voz), aunque también pueden estar integrados en dispositivos físicos (como robots avanzados, vehículos autónomos, drones o aplicaciones del internet de las cosas) (Comisión Europea, 2018).

Lo anterior explica que la Comisión Europea concibe la IA no como una tecnología única ni como una categoría homogénea, sino como un conjunto de tecnologías que combinan datos, algoritmos y capacidad computacional, capaces de integrarse en *software* o en dispositivos físicos y de actuar con distintos grados de autonomía para alcanzar objetivos definidos por seres humanos; por ello, la IA no constituye una noción unitaria ni cerrada, sino un campo tecnológico amplio y dinámico (Comisión Europea, 2020, p. 3). Precisamente debido a este carácter plural y evolutivo, se ha sostenido que la IA no puede

ser regulada en cuanto categoría abstracta, sino únicamente en la medida en que se concreta en productos y servicios específicos con relevancia normativa, razón por la cual cualquier definición jurídica debe ser funcional e instrumental, lo suficientemente flexible para adaptarse al progreso técnico y, al mismo tiempo, suficientemente precisa para otorgar seguridad jurídica, orientándose a los usos y riesgos concretos asociados a determinados contextos sectoriales, en particular cuando se trata de aplicaciones susceptibles de generar riesgos significativos para la seguridad o los derechos fundamentales (Comisión Europea, 2020, pp. 16-18).

Por lo anterior, se ha sostenido que es recomendable emplear la expresión *sistema de IA* para designar cualquier componente, *software* o *hardware* basado en IA (Comisión Europea, 2018, p. 1), pasando, de la definición genérica de IA, a la definición más específica de sistema de IA (SIA) (Navas Navarro, 2022, p. 10).

1.2. Sistemas de inteligencia artificial y su clasificación en función del nivel de riesgos

El Reglamento de la Unión Europea (RIA) —en el cual se basa el proyecto de ley sobre sistemas de IA— define en su artículo 3 numeral 1 al sistema de IA como:

> un sistema basado en una máquina que está diseñado para funcionar con distintos niveles de autonomía y que puede mostrar capacidad de adaptación tras el despliegue, y que, para objetivos explícitos o implícitos, infiere de la información de entrada que recibe la manera de generar resultados de salida, como predicciones, contenidos, recomendaciones o decisiones, que pueden influir en entornos físicos o virtuales.

Por su parte, el proyecto de ley chileno sobre sistemas de IA los define en su artículo 3 numeral 1, como:

> sistema basado en máquinas que, por objetivos explícitos o implícitos infiere, a partir de la entrada que recibe, cómo generar salidas tales como predicciones, contenidos, recomendaciones o decisiones que pueden influir en entornos físicos o virtuales. Los distintos sistemas de IA pueden variar en sus niveles de autonomía y adaptabilidad tras su implementación.

La definición del RIA incorpora la autonomía, la adaptabilidad tras el despliegue y la referencia al diseño del sistema como elementos integrados en la propia definición, con una formulación más extensa y dinámica orientada al ciclo de vida, mientras que la del proyecto de ley chileno reduce la definición al núcleo funcional de inferencia y generación de salidas, desplazando la autonomía y la adaptabilidad a una mención accesoria y variable, sin aludir al diseño ni al momento del despliegue, con una redacción más breve, abstracta y normativamente flexible.

En cualquier caso, se puede concluir que, para estar en presencia de un SIA, tanto en el RIA como en el proyecto chileno, se requiere: a) que se trate de un sistema basado en una máquina; b) que esté diseñado para operar con distintos niveles de autonomía; c) que tenga capacidad de adaptación tras su despliegue; y d) que sea capaz, con miras a objetivos explícitos o implícitos, de generar resultados de salida a partir de la información de entrada recibida, los cuales pueden influir en entornos físicos o virtuales.

Sobre esta base conceptual, el Reglamento de la Unión Europea —que sirve como modelo al Proyecto de Ley chileno sobre sistemas de IA— articula un enfoque por niveles de riesgo que hace recaer las obligaciones jurídicas principales en los sistemas de IA concretos —y no en la IA en abstracto— y, por tanto, en su uso funcional, confirmando que la regulación operativa se construye en torno a los riesgos efectivamente generados por su despliegue (Cf. Martín-Casals, 2022, pp. 119-122). En este contexto, el RIA adopta el riesgo como eje regulatorio fundamental, partiendo de la premisa de que el tipo y el contenido de las normas deben adaptarse a la intensidad y al alcance de los riesgos que los sistemas de IA pueden producir (Mahler, 2025, p. 57, 61), noción que atraviesa todo el texto normativo y sus anexos —invocada en más de 770 ocasiones— y que revela tanto su centralidad como su versatilidad conceptual (Mahler, 2025, p. 61). Este modelo persigue un equilibrio entre la protección de los derechos fundamentales y la promoción de la innovación, calibrando las exigencias normativas en función de la gravedad y probabilidad del daño, con el objetivo de evitar tanto la sobrerregulación como los déficits de control (Mahler, 2025, p. 61).

El enfoque de riesgo en el RIA se presenta, en apariencia, como un modelo ordenado y sencillo gracias a la conocida "pirámide del riesgo"[4]. Esta estructura visual distingue entre riesgo inaceptable, alto, limitado y mínimo, transmitiendo una claridad conceptual que resulta casi intuitiva. Sin embargo, según importante doctrina, la pirámide cumple principalmente un rol comunicativo y heurístico, más que normativo (Mahler, 2025, p. 70), por lo que la configuración de las categorías según el riesgo respondió, en buena medida, a acuerdos alcanzados durante el proceso legislativo, lo que incorpora una carga política y social asociada a las áreas consideradas sensibles (Cf. Mahler, 2025, pp. 63-65). Esta carga política y contingente del enfoque por riesgo se hace especialmente visible cuando el Reglamento se enfrenta a tecnologías que no encajan en una lógica finalista ni sectorial, como ocurre con los modelos y sistemas de inteligencia artificial de uso general.

1.3. Modelos y sistemas de inteligencia artificial de uso general

El artículo 3, número 63, del RIA define al *modelo de IA de uso general* como:

> Un modelo de IA, también uno entrenado con un gran volumen de datos utilizando autosupervisión a gran escala, que presenta un grado considerable de generalidad y es capaz de realizar de manera competente una gran variedad de tareas distintas,

[4] Cf. https://digital-strategy.ec.europa.eu/es/policies/regulatory-framework-ai

> independientemente de la manera en que el modelo se introduzca en el mercado, y que puede integrarse en diversos sistemas o aplicaciones posteriores, excepto los modelos de IA que se utilizan para actividades de investigación, desarrollo o creación de prototipos antes de su introducción en el mercado.

Por su parte, el artículo 3 numeral 66 define el sistema de IA de uso general como: "un sistema de IA basado en un modelo de IA de uso general y que puede servir para diversos fines, tanto para su uso directo como para su integración en otros sistemas de IA".

La distinción entre un modelo de inteligencia artificial de uso general (GPAI Model) y un sistema de inteligencia artificial de uso general (GPAI System) es esencial, pues determina el papel y las responsabilidades de cada uno dentro de la cadena de valor de la inteligencia artificial (Reglamento (UE) 2024/1689, Cdos. 94 y 195). El modelo constituye el componente central de *software* —el motor algorítmico con amplias capacidades—, mientras que el sistema corresponde a la aplicación funcional que se construye sobre dicho modelo. Así, el modelo, por sí solo, no constituye un sistema de IA, ya que requiere otros componentes —como una interfaz de usuario— para operar funcionalmente (Reglamento (UE) 2024/1689, Cdo. 94). Ejemplo de ello son los grandes modelos de IA generativa, capaces de producir texto, imágenes, audio o vídeo, y considerados de alta generalidad cuando superan los mil millones de parámetros y han sido entrenados a gran escala (Reglamento (UE) 2024/1689, Cdos. 98 y 99). En cambio, el sistema de IA de uso general es la materialización funcional del modelo (Cf. Artículo 3.66 del RIA).

Así, la diferencia entre modelo de IA y sistema de IA es clave: el sistema es una categoría más amplia que integra diversos componentes —incluido el modelo— y es el que permite el funcionamiento efectivo y la interacción con usuarios y el entorno, mientras que el modelo es una construcción técnica que genera inferencias o predicciones, pero carece de utilidad autónoma. Así, ChatGPT es un sistema y GPT-3.5 el modelo que lo sustenta. Aunque ambos términos suelen confundirse, la distinción es jurídicamente relevante cuando se imponen obligaciones no solo al sistema final, sino también a sus componentes. En este sentido, el Parlamento optó por regular fases previas de la cadena de valor mediante la noción de modelo fundacional, imponiendo deberes específicos a sus proveedores. Dado que los modelos pueden ser reutilizados por terceros en múltiples aplicaciones, estas obligaciones deben ser proporcionadas y acompañarse de deberes de transparencia para evitar asimetrías de información y asegurar una asignación equilibrada de responsabilidades (Fernández-Llorca et al., 2025, pp. 881-883).

Por tanto, el modelo GPAI representa el componente genérico y de propósito amplio —el qué se entrena y utiliza—, mientras que el sistema GPAI constituye la aplicación práctica y operativa —el cómo se despliega e interactúa con el entorno— (Parlamento Europeo, 2024, Cdo. 94) (artículos 3.1, 3.63 y 3.66 del RIA). En otras palabras, el modelo es el "qué" se entrena y se pone a disposición en el mercado; y, el sistema, es el "cómo" se despliega ese modelo en un entorno especifico.

En este contexto, conviene precisar que el proyecto de ley chileno, a diferencia del Reglamento de Inteligencia Artificial de la Unión Europea, no contempla ni regula los modelos ni los sistemas de IA de uso general (Cf. Amunátegui Perelló, 2025). Esta omisión se explica porque el proyecto chileno reproduce en gran medida la versión inicial del RIA: la Propuesta de 2021, que aún no incorporaba esta categoría regulatoria. Los modelos y sistemas de uso general fueron introducidos recién en la fase final del proceso legislativo europeo, razón por la cual no se reflejan en el texto chileno.

Desde esta distinción entre modelos y sistemas de IA de uso general se comprende adecuadamente la posición normativa de la inteligencia artificial generativa. Así, como sostiene el RIA:

> Los grandes modelos de IA generativa son un ejemplo típico de un modelo de IA de uso general, ya que permiten la generación flexible de contenidos, por ejemplo, en formato de texto, audio, imágenes o vídeo, que pueden adaptarse fácilmente a una amplia gama de tareas diferenciadas. (RIA, Cdo. 99)[5]

1.4. Inteligencia artificial generativa (GenAI) propiamente tal

La GenIA es una tecnología diseñada para crear contenidos como texto, imágenes, audio (incluida música) y vídeo. Estos sistemas utilizan aprendizaje profundo y redes neuronales para detectar patrones y predecir respuestas a partir de cálculos matemáticos y estadísticos, generando contenidos similares a los datos con los que han sido entrenados, en función de las instrucciones del usuario. La IA generativa puede habilitar nuevas formas de comunicación y expresión y no posee intenciones propias, como engañar. No obstante, puede ser utilizada por las personas no solo para maximizar beneficios económicos, sino también para facilitar o crear campañas de influencia engañosas, que resultan más persuasivas y difíciles de detectar que en el pasado. Así, la GenAI reduce significativamente los costos y barreras para la manipulación e injerencia informativa, al facilitar y automatizar la creación de contenidos, al tiempo que presenta vulnerabilidades estructurales a lo largo de todo su ciclo de vida, como sesgos o contaminación intencional de los datos de entrenamiento, manipulación de algoritmos y uso malicioso de instrucciones, lo que favorece la generación y amplificación de desinformación, *deepfakes*, discursos de odio, extremismo y violencia de género en línea (Bentzen, 2025, pp. 1-2).

Aunque el aprendizaje generativo es un concepto clásico del *machine learning*, el término ha adquirido una relevancia renovada con la difusión masiva de aplicaciones como ChatGPT o Midjourney. En el plano normativo, la generación de contenidos ya se encuentra comprendida en las definiciones de sistema de IA de la Comisión y el Consejo

[5] Este considerando se refiere exclusivamente a los modelos de IA porque la IA generativa es, desde el punto de vista técnico y regulatorio, una característica del modelo y no del sistema. Lo que hace que una IA sea "generativa" es su capacidad intrínseca de producir contenidos nuevos, capacidad que reside en el modelo entrenado a gran escala y no en la aplicación concreta en la que posteriormente se despliega.

Europeo, que además imponen obligaciones específicas de transparencia frente a los *deepfakes* (Fernández-Llorca et al., 2025, pp. 885-886).

En este contexto, la IA generativa (GenAI) constituye una subdisciplina central de la IA orientada a la creación de nuevos datos a partir de patrones existentes, con un impacto particularmente intenso en la comunicación, el periodismo y las redes sociales. Basada en técnicas de aprendizaje profundo y procesamiento del lenguaje natural, permite la generación automatizada y multimodal de contenidos, transformando los procesos de producción y circulación de la información y planteando, al mismo tiempo, relevantes desafíos éticos, regulatorios y comunicacionales que exigen marcos normativos adecuados para un desarrollo responsable (García de Blanes et al., 2025, pp. 3-5).

Desde la perspectiva del Reglamento de IA, la IA generativa, en cuanto gran modelo entrenado a gran escala, se inserta en la categoría de modelos de IA de uso general, en la medida en que permite una generación flexible de contenidos —texto, audio, imágenes o video— adaptable a una amplia gama de tareas diferenciadas (Reglamento (UE) 2024/1689, 2024, Cdo. 99). Estos modelos se caracterizan por su elevada generalidad funcional, su entrenamiento con grandes volúmenes de datos y su comercialización en distintas modalidades (Reglamento (UE) 2024/1689, Cdo. 97). Cuando tales modelos se integran con otros componentes técnicos y organizativos para ofrecer funcionalidades concretas al usuario, pasan a constituir sistemas de IA de uso general, conforme al artículo 3 numeral 66 del RIA.

En consecuencia, la IA generativa no constituye una categoría autónoma: es modelo de IA de uso general cuando alude al núcleo algorítmico entrenado a gran escala con capacidades generativas amplias (artículo 3, numeral 63), y es sistema de IA de uso general cuando dicho modelo se despliega en una aplicación concreta orientada a un uso específico (artículo 3, numeral 66). La distinción es decisiva, pues el modelo representa el componente general y reutilizable, mientras que el sistema corresponde a su concreción funcional, normalmente el punto en que se activan los deberes de uso, supervisión y responsabilidad a lo largo de la cadena de valor del RIA.

La relevancia actual de la GenAI radica, finalmente, en que ya se encuentra ampliamente implementada en ámbitos como la educación, la salud, la investigación científica, la administración pública y la industria, planteando problemas significativos en materia de sesgos, transparencia, privacidad, impacto laboral y desinformación (Comisión Europea, 2025), entre los que destaca el problema de los "*deepfake*", objeto de este trabajo.

2. Imitaciones no consentidas mediante inteligencia artificial (*deepfake*)

2.1. Concepto de *deepfake*

El término *deepfake* combina "*deep learning*" (aprendizaje profundo) y "*fake*" (falso). El RIA define los *deepfakes* como "un sistema de IA que genera o manipula contenidos

visuales y de audio", que "se asemeja de manera apreciable a personas, objetos, lugares u otras entidades o acontecimientos existentes y que falsamente parecería a una persona ser auténtico o veraz" (RIA, Cdo 134).

La tecnología *deepfake* es cada vez más accesible y puede utilizarse con fines de entretenimiento o comunicación. Sin embargo, el uso malicioso y fraudulento de los *deepfakes* puede plantear riesgos para los procesos democráticos, las instituciones, así como para las empresas, los grupos vulnerables y las personas individuales (Bentzen, 2025, p. 3).

Así, se trata de contenidos sintéticos de audio, video o imagen generados mediante técnicas de IA que reproducen de forma altamente realista a personas, situaciones o acontecimientos que nunca ocurrieron, alcanzando un grado de verosimilitud que dificulta distinguir entre contenido auténtico y manipulado y favorece la aceptación de información falsa; su especial relevancia en la propagación de la desinformación se explica por la apariencia de autenticidad que presentan, por su elevada viralidad en entornos digitales debido a su impacto y rapidez de difusión, por su capacidad de generar falsas representaciones mediante la suplantación de identidades o la alteración de declaraciones con efectos potencialmente graves sobre reputaciones y confianza pública, y por su uso como herramienta de desinformación dirigida, al permitir reforzar narrativas específicas, explotar sesgos preexistentes o intensificar tensiones sociales (Dhiman, 2023, pp. 1-2).

En este contexto, se ha sostenido que se trata de contenido digital manipulado o generado mediante técnicas avanzadas de inteligencia artificial, especialmente aprendizaje profundo, para simular eventos o comportamientos que nunca ocurrieron (European Data Protection Supervisor, 2023; Patil et al., 2023, p. 2). Asimismo, se ha definido como "el uso de técnicas de aprendizaje profundo para fabricar contenidos mediáticos engañosos" (Ramos-Zaga, 2024, p. 365) o como "aquellos videos, imágenes o audios que son generados por parte de la IA y que tratan de imitar la apariencia y el sonido de una persona" (Aragüez, 2024, p. 49).

Así, el elemento común es la caracterización del *deepfake* como contenido sintético altamente realista, generado mediante aprendizaje profundo, diseñado para simular la realidad de manera engañosa, con capacidad de inducir a error y producir efectos sociales relevantes.

De esta forma, los *deepfakes* tienen un impacto social relevante al debilitar la confianza en la información, los medios y las instituciones, al dificultar la distinción entre contenidos auténticos y manipulados. Facilitan la desinformación, la manipulación de la opinión pública y la difamación, al tiempo que generan graves riesgos para la privacidad, el consentimiento y la reputación de las personas, incluyendo posibles usos con fines de extorsión. Estas amenazas se ven intensificadas por las dificultades para su detección y por la insuficiencia de los marcos jurídicos actuales, lo que compromete la integridad del ecosistema informativo y plantea complejos desafíos éticos y legales (Dhiman, 2023, pp. 2-3).

Los *deepfakes* pueden clasificarse en distintas categorías según el tipo de contenido manipulado. Entre las más comunes se encuentran el intercambio de rostros, que sustituye la cara de una persona por la de otra en imágenes o videos; la clonación de voz, que reproduce artificialmente la voz de un individuo; y los *deepfakes* textuales, que imitan el estilo y contenido de la escritura de una persona mediante modelos de procesamiento del lenguaje natural. A ello se suman los contenidos audiovisuales sintéticos, que crean o alteran escenas y acontecimientos inexistentes, la manipulación de gestos y comportamientos, que modifica movimientos corporales para generar impresiones engañosas, y los *deepfakes* multimodales, que combinan audio, video y texto, incrementando su realismo y dificultad de detección (Dhiman, 2023, pp. 3-4).

La literatura técnica clasifica los *deepfakes* en dos grandes grupos: visuales (*face-swap, lip-syncing, face-reenactment o puppet-master*, síntesis facial completa y manipulación de atributos) y auditivos (*text-to-speech* y *voice conversion*). Cada categoría cuenta con métodos avanzados de generación basados en *autoencoders*[6], GANs, modelos 3D y técnicas de transferencia de expresión, que permiten producir rostros intercambiados, sincronizar labios con cualquier audio, reenactar movimientos faciales o generar caras inexistentes con gran realismo. Paralelamente, la investigación en detección emplea tanto técnicas tradicionales (*landmarks*, óptica, señales fisiológicas, co-ocurrencias) como modelos profundos (CNN, RNN, LSTM, XceptionNet, redes siamesas) para identificar inconsistencias temporales, artefactos de renderizado, desalineación audio-video o huellas estadísticas propias de imágenes sintéticas. A pesar de los avances, los detectores siguen enfrentando limitaciones frente a videos comprimidos, muestras no vistas, grandes variaciones de pose y manipulaciones cada vez más sofisticadas, lo que mantiene un desafío creciente en la verificación de autenticidad audiovisual (Cf. Masood et al., 2023, p. 7-20).

Las manipulaciones de atributos faciales y las voces sintéticas han avanzado gracias a arquitecturas GAN y modelos neuronales que permiten editar rasgos como edad, género, cabello o accesorios, así como clonar voces mediante TTS y conversión de voz, aunque con limitaciones en preservación de identidad, calidad bajo ruido y rendimiento en modificaciones complejas. Las técnicas de detección, basadas en *handcrafted features* y en *deep learning*, logran altos niveles de precisión aprovechando huellas estadísticas, artefactos visuales, inconsistencias temporales y desalineaciones fisiológicas, aunque su desempeño cae ante muestras comprimidas, escenarios reales, eliminación de huellas GAN y ataques adversarios. Persisten desafíos fundamentales tanto en la generación como en la detección: necesidad de modelos más generalizables, problemas de identidad filtrada o distorsionada, dependencia del entrenamiento pareado, sensibilidad a pose, iluminación y oclusiones, falta de realismo en audio, baja calidad y sesgos en los *datasets*, ausencia de explicabilidad, dificultad para detectar manipulación multimodal y para operar en contenido

[6] Dos pares *encoder-decoder* comparten un *encoder* común para aprender rasgos faciales e intercambiar identidades entre imágenes El sistema requiere cientos o miles de imágenes del sujeto para generar resultados convincentes (Patil et al., 2023, p. 3).

alterado por redes sociales. En síntesis, la sofisticación creciente de los *deepfakes* y la aparición de técnicas diseñadas para evadir detectores plantean un escenario donde las defensas actuales resultan frágiles frente a manipulaciones cada vez más realistas y difíciles de rastrear (Cf. Masood et al., 2023, p. 20-34).

Los *deepfakes* plantean importantes implicancias jurídicas, ya que los marcos normativos vigentes suelen resultar insuficientes para abordar su creación y difusión. En particular, generan desafíos en materia de protección de la privacidad y de los datos personales, al permitir el uso no consentido de la imagen o identidad de las personas; en el ámbito de la propiedad intelectual, por la utilización indebida de obras protegidas o de la apariencia de terceros; y en relación con la difamación, al facilitar la atribución de conductas o declaraciones falsas con daño reputacional. Asimismo, su potencial uso delictivo en fraudes, suplantación de identidad u otras actividades criminales exige una actualización de las normas penales, mientras que a nivel regulatorio se discuten mecanismos preventivos, como obligaciones de etiquetado de contenidos manipulados y deberes reforzados de detección y eliminación para las plataformas digitales (Dhiman, 2023, pp. 5-6).

Los *deepfakes* se han convertido en una herramienta poderosa para la desinformación, al permitir la creación de videos y audios falsos extremadamente convincentes que pueden dañar reputaciones, extorsionar, manipular procesos electorales, generar conflictos y fortalecer narrativas políticas o conspirativas. Su impacto se amplifica por diversos actores que difunden contenido manipulado: *trolls* independientes o contratados, *bots* automatizados, teorías conspirativas, medios hiperpartidistas, políticos que buscan desacreditar adversarios y gobiernos extranjeros que emplean propaganda digital. En conjunto, estos actores utilizan los *deepfakes* para intensificar la difusión de información falsa, afectar la opinión pública y aumentar la inestabilidad social y geopolítica (Masood et al., 2023, pp. 3-4).

La proliferación de *deepfakes* plantea riesgos significativos: difusión de desinformación, afectación de la opinión pública, generación de contenido ofensivo, fraudes financieros y suplantación de identidad con posibles accesos indebidos a servicios o información. Para enfrentarlos, se han desarrollado técnicas de detección basadas en señales fisiológicas o visuales (parpadeo, movimiento, reflejos o variaciones de la piel) y en la identificación de artefactos característicos del contenido sintetizado. No obstante, los detectores actuales presentan limitaciones por depender de firmas conocidas, por la escasez y ruido de los datos de entrenamiento y por su insuficiente capacidad para identificar manipulaciones novedosas (European Data Protection Supervisor, 2023).

En protección de datos, la detección de *deepfakes* tiene impactos positivos: permite prevenir la circulación de contenido falso que afecte la privacidad o reputación de las personas; dificulta ataques de suplantación de identidad utilizados en phishing y otras formas de ingeniería social; y posibilita validar la autenticidad de información crítica en sectores regulados, mejorando la fiabilidad de los datos (European Data Protection Supervisor, 2023).

Pero también existen impactos negativos relevantes. Los conjuntos de datos usados para entrenar detectores suelen ser poco diversos y generan sesgos que afectan la equidad, discriminando a personas según rasgos físicos, origen étnico o género. Además, los modelos de detección, especialmente los de aprendizaje profundo, carecen de transparencia y explicabilidad, lo que dificulta su utilización en contextos periodísticos, probatorios o legales. Finalmente, su precisión es limitada: tratan el problema como una clasificación binaria, no consideran otras formas legítimas de edición y fallan ante el lavado de medios presente en redes sociales, aumentando los falsos negativos (European Data Protection Supervisor, 2023).

En este contexto, la detección y mitigación de los *deepfakes* presenta importantes desafíos debido a su creciente sofisticación tecnológica, lo que ha impulsado el desarrollo de diversas estrategias complementarias. Entre las principales se encuentran el análisis forense, orientado a identificar inconsistencias técnicas en imágenes, videos o audios; el uso de marcas de agua digitales y sistemas de autenticación, que permiten verificar el origen y la integridad del contenido; y la aplicación de algoritmos de aprendizaje automático, entrenados para detectar patrones propios de contenidos manipulados. A ello se suman tecnologías de reconocimiento facial y de voz, el análisis de comportamientos anómalos en la difusión del contenido, y soluciones basadas en *blockchain* para garantizar trazabilidad y autenticidad. Estas herramientas técnicas se complementan con iniciativas colaborativas entre actores públicos y privados y con el desarrollo de políticas y marcos normativos, orientados a prevenir, detectar y sancionar el uso malicioso de *deepfakes* (Dhiman, 2023, pp. 4-5).

No obstante, lo cierto es que la expansión de los *deepfakes* plantea relevantes implicancias éticas y jurídicas, al afectar directamente la privacidad, la dignidad y los derechos fundamentales de las personas, así como la confianza social en la información. Desde una perspectiva ética, destacan los problemas asociados al uso no consentido de la imagen y la identidad personal, la manipulación y tergiversación de individuos con potencial daño reputacional y emocional, y la erosión de la verdad y la confianza pública al difuminarse la frontera entre realidad y ficción, con efectos adversos sobre la cohesión social y el debate democrático. En el plano jurídico, los *deepfakes* revelan las limitaciones de los marcos normativos vigentes en materia de protección de la privacidad, propiedad intelectual y difamación, además de su potencial utilización para fines delictivos como el fraude o la suplantación de identidad, lo que ha impulsado la discusión sobre nuevas estrategias regulatorias, incluyendo obligaciones de etiquetado de contenidos manipulados y deberes reforzados de detección y retirada por parte de las plataformas digitales (Dhiman, 2023, pp. 5-6).

La Unión Europea ha respondido a la manipulación informativa asociada a la IA mediante un marco normativo integrado que, a través del Reglamento de Inteligencia Artificial de 2024, prohíbe técnicas subliminales o manipulativas, impone obligaciones de transparencia y etiquetado de contenidos artificiales como los *deepfakes* y califica como

de alto riesgo los sistemas destinados a influir en procesos electorales; del Reglamento de Servicios Digitales, que obliga a las grandes plataformas a evaluar y mitigar riesgos sistémicos, incluyendo el etiquetado de contenidos manipulados y la cooperación con denunciantes de confianza; del Código de Conducta sobre Desinformación, que establece compromisos específicos en materia de inteligencia artificial generativa y transparencia; de la Directiva (Unión Europea) 2024/1385, que criminaliza los abusos sexuales basados en imágenes, incluso cuando se generan o manipulan mediante IA; y del Reglamento Europeo sobre la Libertad de los Medios, que limita la protección reforzada a contenidos sometidos a control editorial humano efectivo (Cf. Bentzen, 2025, pp. 9-12).

En el caso chileno, resultan especialmente relevantes tres proyectos de ley en tramitación: el Boletín N°16.281-19, de 7 de mayo de 2024, que regula los sistemas de inteligencia artificial; el Boletín N°17.307-07, de 17 de diciembre de 2024, que modifica el Código Penal con el objeto de tipificar la generación y difusión de imágenes o hechos de carácter privado o íntimo creados mediante herramientas de inteligencia artificial; y, finalmente, el Boletín N°17.795-19, de 21 de agosto de 2025, que establece protección frente al uso no consentido de tecnologías de inteligencia artificial que imitan la imagen, el cuerpo o la voz de las personas, siendo este último el de particular relevancia para los fines del presente trabajo.

2.2. Proyectos de ley chilenos

2.2.1. Boletín N°16.281-19 (7 de mayo de 2024): Proyecto de ley "Que regula los sistemas de inteligencia artificial"

El proyecto de ley chileno no regula los *deepfakes* como una categoría normativa autónoma ni emplea expresamente dicho concepto. Sin embargo, determinadas manifestaciones de *deepfake* pueden quedar comprendidas de manera indirecta dentro de su ámbito de aplicación, en función del uso concreto del sistema, del nivel de riesgo asociado y de la eventual afectación de derechos fundamentales. Esta aproximación fragmentaria contrasta con la relevancia práctica y jurídica que han adquirido los contenidos sintéticos hiperrealistas en el espacio digital.

En efecto, el articulado del proyecto no contiene ninguna referencia al término *deepfake* ni incorpora una definición funcional equivalente, como contenido sintético, imitaciones audiovisuales realistas o suplantación digital. Tampoco establece un deber general y explícito de rotulación o advertencia respecto del contenido generado o manipulado mediante inteligencia artificial, a diferencia de lo que ocurre en el artículo 52 del Reglamento (UE) 2024/1689. Esta omisión constituye una laguna normativa significativa, en tanto priva al ordenamiento de un criterio claro de identificación y tratamiento jurídico de este tipo de contenidos.

Pese a ello, algunos *deepfakes* podrían quedar alcanzados por el proyecto a través de categorías generales de riesgo (artículo 5), cuya aplicación es necesariamente casuística. En particular, aquellos usos que afecten gravemente la honra, la integridad personal o

el libre desarrollo de la sexualidad, especialmente cuando se trate de niños, niñas o adolescentes, o que exploten vulnerabilidades y generen formas intensas de manipulación (artículo 6, letras A-G), podrían ser calificados como sistemas de inteligencia artificial de riesgo inaceptable y, por ende, quedar prohibidos. En esta categoría se inscribirían, por ejemplo, los *deepfakes* sexuales no consentidos o las suplantaciones identitarias especialmente lesivas.

Otros *deepfakes* de menor intensidad dañosa, como aquellos de carácter recreativo, artístico o paródico, podrían eventualmente encuadrarse dentro de los sistemas de riesgo limitado, quedando sujetos a deberes genéricos de transparencia consistentes en informar que la persona interactúa con un sistema de IA (artículo 12). No obstante, el proyecto no prevé obligaciones específicas de advertencia audiovisual ni reglas adaptadas a la naturaleza engañosa propia de los *deepfakes*, lo que reduce la eficacia de estos deberes[7].

Desde la perspectiva de la responsabilidad civil, el proyecto permite accionar por culpa cuando el uso de un sistema de IA cause un daño, conforme a sus disposiciones generales (artículo 28). Sin embargo, no contempla un régimen especial que atienda a las particularidades técnicas, probatorias y causales de los *deepfakes*, lo que dificulta la imputación de responsabilidad y la tutela efectiva de las víctimas en este ámbito[8].

En conjunto, el proyecto no aborda los *deepfakes* como un fenómeno jurídico diferenciado, sino que los absorbe de manera insuficiente y dispersa dentro de categorías amplias de riesgo y afectación de derechos fundamentales. Este enfoque genera una relevante incerteza normativa, especialmente frente a *deepfakes* que no queden comprendidos en las prácticas de riesgo inaceptable, como aquellos no sexuales, de carácter político, informativo o satírico, resultando insuficiente a la luz del estándar europeo vigente, lo que explica la creciente proliferación de iniciativas legislativas específicas sobre imitaciones no consentidas y contenidos sintéticos en el ordenamiento chileno.

2.2.2. Boletín N°17.307-07 (17 de diciembre de 2024): Proyecto de Ley "Que modifica el Código Penal, con el objeto de tipificar la generación y difusión de imágenes o hechos de carácter privado o íntimo, creados con herramientas de inteligencia artificial"

Este proyecto, estuvo destinado a modificar el Código Penal para tipificar y sancionar la generación y difusión, mediante sistemas de inteligencia artificial —en particular, *deepfakes*—, de imágenes, videos u otros contenidos de carácter íntimo sin consentimiento, con el objeto de proteger la privacidad y la dignidad de las personas frente al incremento del abuso y la violencia digital[9].

[7] El artículo 12 del proyecto simplemente sostiene que los sistemas de IA de riesgo limitado deben informar clara y oportunamente a las personas cuando interactúan con IA, salvo que sea evidente por el contexto

[8] Sobre las críticas al tratamiento de la responsabilidad civil en el proyecto de ley chileno sobre IA, véase: Diez Ringele, 2024, pp. 221-244.

[9] Senado de la República de Chile, Proyecto de ley que modifica el Código Penal, con el objeto de tipificar la generación y difusión de imágenes o hechos de carácter privado o íntimo, creados con herramientas de inteligencia artificial, Boletín N°17.307-07.

En síntesis, el proyecto pretende modificar los artículos 161-A y 161-C del Código Penal para sancionar penalmente la captación, generación —incluida la realizada mediante inteligencia artificial— y difusión no consentida de comunicaciones, documentos, imágenes o hechos de carácter privado, así como de imágenes íntimas con significación sexual, tanto en espacios privados como públicos, agravando las penas cuando una misma persona obtiene y difunde el material, e incorporando expresamente la creación de contenidos falsos o manipulados (*deepfakes*) como una forma relevante de vulneración de la vida privada y la intimidad personal[10].

2.2.3. Boletín N°17.795-19 (21 de agosto de 2025): Proyecto de Ley "Que establece protección a las personas frente al uso no consentido de tecnologías de inteligencia artificial que imitan su imagen, cuerpo o voz"

Según los antecedentes del mismo, busca proteger a las personas frente al uso no consentido de tecnologías de inteligencia artificial que imitan su imagen, cuerpo o voz (*deepfakes*), ante el vacío normativo existente en Chile y los riesgos que estas prácticas suponen para la privacidad, la identidad y la dignidad personal. La iniciativa reconoce la imagen, el cuerpo y la voz como expresiones de la identidad, establece un derecho exclusivo de control sobre sus imitaciones digitales, exige consentimiento previo y revocable, impone obligaciones a las plataformas, contempla sanciones y mecanismos de reparación, y se articula con el marco constitucional y de protección de datos personales vigente[11].

Este proyecto de 2025 se estructura en cuatro títulos y 12 artículos. El primer título trata los objetos y definiciones; el segundo, el derecho a la integridad digital (II); el tercero, las responsabilidades y sanciones; y, finalmente, la coordinación del proyecto con otras normas.

2.3. Referencia comparada internacional: el artículo 35.1.k del Reglamento de Servicios Digitales de la Unión Europea y la Take It Down Act de los Estados Unidos

El análisis de los proyectos de ley chilenos no puede desvincularse del contexto normativo internacional, en el que dos instrumentos recientes —uno europeo y uno estadounidense— ofrecen modelos de regulación que iluminan, por comparación, tanto los aciertos como las insuficiencias del Boletín N°17.795-19.

2.3.1. El artículo 35.1.k del Reglamento de Servicios Digitales (DSA) de la Unión Europea

El Reglamento (UE) 2022/2065, sobre el mercado único de servicios digitales —Digital Services Act (DSA)—, no regula los *deepfakes* como fenómeno autónomo, sino que los aborda a través de un régimen de gestión de riesgos sistémicos aplicable a las plataformas

[10] Esta propuesta parece seguir el proyecto español, que opta por un enfoque penal, proponiendo incorporar el artículo 173 bis al Código Penal para sancionar la difusión de *deepfakes* sexuales o gravemente vejatorios sin consentimiento y con ánimo de dañar la integridad moral, con agravantes cuando las víctimas son menores, personas con discapacidad o cuando existe difusión masiva.

[11] https://www.camara.cl/legislacion/proyectosdeley/tramitacion.aspx?prmID=18447&prmBOLETIN=17795-19

de muy gran tamaño (Very Large Online Platforms, VLOPs), esto es, aquellas que superan los 45 millones de usuarios activos mensuales en la Unión Europea

Su artículo 35, titulado "Reducción de riesgos", establece en su numeral 1 que:

> Los prestadores de plataformas en línea de muy gran tamaño y de motores de búsqueda en línea de muy gran tamaño aplicarán medidas de reducción de riesgos razonables, proporcionadas y efectivas, adaptadas a los riesgos sistémicos específicos detectados de conformidad con el artículo 34, teniendo especialmente en cuenta las consecuencias de dichas medidas sobre los derechos fundamentales. Dichas medidas podrán incluir, cuando proceda: k) garantizar que un elemento de información, ya se trate de imagen, audio o vídeo generado o manipulado que se asemeja notablemente a personas, objetos, lugares u otras entidades o sucesos existentes y que puede inducir erróneamente a una persona a pensar que son auténticos o verídicos, se distinga mediante indicaciones destacadas cuando se presente en sus interfaces en línea y, además, proporcionar una funcionalidad fácil de utilizar que permita a los destinatarios del servicio señalar dicha información.

Esta disposición es significativa por dos razones. Primera, porque adopta una lógica de transparencia antes que supresión, es decir, la obligación principal de la plataforma no es retirar el *deepfake* lícito sino etiquetarlo, lo que preserva la libertad de expresión mientras informa al usuario. Segunda, porque la obligación recae sobre la plataforma como parte de su gestión proactiva de riesgos —no solo como reacción a una denuncia individual—, lo que implica una diferencia estructural relevante respecto del modelo reactivo de 72 horas que contempla el artículo 9 del Boletín N°17.795-19. El incumplimiento de estas obligaciones puede ser sancionado con multas de hasta el 6% del volumen de negocios anual mundial del infractor, estándar que contrasta también con el régimen sancionatorio del proyecto chileno, que fija una multa de 100 a 1.000 UTM sin criterios de graduación vinculados a la capacidad económica del responsable.

2.3.2. La Take It Down Act de los Estados Unidos

En el ordenamiento estadounidense, la respuesta legislativa federal más relevante es la Tools to Address Known Exploitation by Immobilizing Technological Deepfakes on Websites and Networks Act —conocida como Take It Down Act—, promulgada el 19 de mayo de 2025 (Take It Down Act, 2025, §146). La ley prohíbe que cualquier persona publique a sabiendas, sin consentimiento, representaciones visuales íntimas de menores o adultos no consintientes, así como cualquier *deepfake* —sea o no de carácter íntimo— cuya publicación esté destinada a causar daño Skadden. Adicionalmente, impone a las plataformas —sitios web públicos y aplicaciones móviles— la obligación de establecer procedimientos de notificación y retiro (*notice and takedown*), debiendo eliminar los contenidos denunciados en un plazo máximo de 48 horas.

La Take It Down Act es relevante para el análisis del proyecto chileno por al menos dos razones. Por un lado, es la primera ley federal estadounidense que regula directamente el uso de IA en contextos potencialmente dañosos para personas individuales, lo que la convierte en un punto de referencia ineludible en el derecho comparado sobre *deepfakes*. Por otro, su plazo de retiro de 48 horas para las plataformas es más exigente que el de 72 horas que establece el artículo 9 del Boletín N°17.795-19, diferencia que no parece justificada en términos de política regulatoria, dado que el daño reputacional, psicológico o patrimonial que generan estos contenidos se produce en las primeras horas de circulación.

2.3.3. Valoración comparada

El cotejo con el DSA y la Take It Down Act revela que el Boletín N°17.795-19, aunque avanza en la dirección correcta, presenta un diseño de obligaciones de plataformas que es, a la vez, menos preventivo que el europeo —por carecer de una lógica de gestión proactiva de riesgos y etiquetado— y menos exigente en sus plazos de retiro que el estadounidense. Esta brecha no es trivial: en un ecosistema digital global donde las mismas plataformas operan simultáneamente bajo el DSA, la Take It Down Act y —eventualmente— la ley chilena, el estándar más débil tiende a convertirse en la norma de facto para los usuarios que quedan fuera de las jurisdicciones más protegidas (Roberts, 2025, pp. 1-10).

3. Objetos, ámbito de aplicación y definiciones del proyecto de ley (Boletín 17.795-19)

3.1. Objeto

El artículo 1 del Proyecto de Ley dispone, respecto al objeto de la ley, que:

> La presente ley tiene por objeto proteger la identidad e integridad de las personas frente al uso no consentido de tecnologías de inteligencia artificial (en adelante IA) para generar, difundir o almacenar imitaciones digitales realistas de su imagen, cuerpo o voz.

3.1.1. Qué debe entenderse por uso: la diferencia con la comercialización y puesta en servicio

El primer aspecto relevante sobre el objeto del proyecto de ley es que se refiere al "uso" de la inteligencia artificial. Sobre esto, cabe destacar que el "uso" no se encuentra definido en el proyecto de ley chileno ni en el RIA. No obstante, conforme a las recientes directrices emitidas por la Comisión Europea, abarca cualquier forma de utilización o despliegue del sistema en cualquiera de las etapas de su ciclo de vida posteriores a su comercialización o puesta en servicio. Ello incluye su integración en los servicios y procesos del o los usuarios, incluso como parte de sistemas, procesos o infraestructuras más complejas (Comisión Europea, 2025, p. 5). Por tanto, el problema de la disposición es que presenta una laguna al no comprender la comercialización ni la puesta en servicio.

El proyecto chileno que regula la IA no define lo que es la comercialización, lo que sí hace el RIA, en el artículo 3 numeral 10, como: "el suministro de un sistema de IA o de un modelo de IA de uso general para su distribución o utilización en el mercado de la Unión en el transcurso de una actividad comercial, previo pago o gratuitamente".

Por su parte, la puesta en servicio sí es definida por el proyecto chileno, la que conforme al artículo 3, numeral 12, la define como "el suministro de un sistema de IA para su primer uso directamente por parte del implementador o para uso propio en el mercado nacional, a título gratuito u oneroso, de acuerdo con su finalidad prevista".

Por su parte, el RIA, en el artículo 3 apartado 9, la define como "la primera comercialización en el mercado de la Unión de un sistema de IA o de un modelo de IA de uso general".

Así, abarca tanto la venta como la oferta gratuita, y puede realizarse a través de cualquier medio (API, nube, descargas, copias físicas o integradas en productos) (Comisión Europea, 2025, p. 4). Por ejemplo, un sistema de identificación biométrica remota (RBI) desarrollado fuera de la UE se considera "introducido en el mercado" cuando se ofrece a usuarios en uno o más Estados miembros, incluso online (Comisión Europea, 2025, p. 4).

La finalidad prevista, conforme al artículo 3 numeral 12 del RIA es "el uso para el que un proveedor concibe un sistema de IA, incluidos el contexto y las condiciones de uso concretos, según la información facilitada por el proveedor en las instrucciones de uso, los materiales y las declaraciones de promoción y venta, y la documentación técnica". Por su parte, el artículo 3 numeral 12 define la puesta en servicio como: "el suministro de un sistema de IA para su primer uso directamente por parte del implementador o para uso propio en el mercado nacional, a título gratuito u oneroso, de acuerdo con su finalidad prevista".

En otras palabras, la puesta en servicio es el momento en que el sistema de IA empieza a funcionar efectivamente (primer uso) para cumplir el propósito que motivó su creación (finalidad prevista; por ejemplo, vigilar, analizar, clasificar, decidir, etc.).

Por tanto, el problema se presenta porque el proyecto de ley chileno solo abarca el uso, mas no comercialización o puesta en servicio, lo que deja una laguna para proteger la identidad e integridad de las personas frente a la comercialización o puesta en servicio de IA no consentida.

3.1.2. Tecnologías de inteligencia artificial

Como se sostuvo anteriormente, la IA es un concepto abierto, impreciso y evolutivo, sin definición consensuada. Su carácter borroso, polifacético y dinámico impide regularla como categoría abstracta, por lo que el enfoque jurídico se desplaza hacia los sistemas de IA, entendidos funcionalmente según sus usos y riesgos concretos. La Comisión Europea concibe la IA como un conjunto de tecnologías basadas en datos y algoritmos, integrables en *software* o dispositivos físicos, capaces de actuar con distintos grados de autonomía para alcanzar objetivos definidos por humanos. Por lo anterior, el Reglamento de IA de la UE y el proyecto de ley chileno definen al sistema de IA —y no a la IA— como un sistema basado en máquinas que infiere, a partir de datos de entrada, resultados de salida que

influyen en entornos físicos o virtuales. Sobre esta base, el RIA y el proyecto chileno sobre IA estructuran su regulación en un enfoque por niveles de riesgo, haciendo recaer las obligaciones jurídicas en los sistemas concretos y en su uso, no en la IA en abstracto, lo que hace que la precisión del objeto de este proyecto resulte imprecisa e insuficiente.

3.2. Ámbito de aplicación

En cuanto al ámbito de aplicación, el artículo 2 del Proyecto de Ley, reza:

> Esta ley se aplica a toda persona natural, nacional o extranjera, viva o fallecida, cuya imagen, cuerpo o voz sean utilizados mediante IA para crear contenidos audiovisuales o sonoros realistas sin autorización. Se aplica también a plataformas que difundan, reproduzcan o mantengan disponible dicho contenido en territorio nacional o accesible desde Chile.

En este sentido, quedan fuera del ámbito de aplicación de esta ley el nombre, los datos personales no audiovisuales, el estilo o forma de pensar, los gestos o conductas genéricas, las imitaciones no realistas o ficticias, los personajes inspirados, pero no identificables, y la honra o reputación como bienes autónomos; la norma solo protege la reproducción realista y reconocible de la imagen de una persona —en términos comparables al derecho a la propia imagen—, así como de su cuerpo o voz, mediante sistemas de inteligencia artificial.

Sin embargo, en Chile, esos elementos —excluidos por la ley— encuentran tutela principalmente en la Constitución (artículo 19 n.º 4: vida privada, honra e identidad personal) y, de manera autónoma, en el derecho a la protección de los datos personales, reconocido constitucionalmente y desarrollado por la Ley 21.719, que comenzará a regir el 1 de diciembre de 2026. Este último permite extender la protección a atributos como la imagen, la voz o los datos biométricos, en la medida en que constituyen datos personales —especialmente sensibles— cuando identifican o hacen identificable a una persona.

Así, la nueva Ley de Protección de Datos establece en su artículo 8 bis. Inciso 1:

> Decisiones individuales automatizadas, incluida la elaboración de perfiles. El titular de datos tiene derecho a oponerse y a no ser objeto de decisiones basadas en el tratamiento automatizado de sus datos personales, incluida la elaboración de perfiles, que produzca efectos jurídicos en él o le afecte significativamente.

Por su parte, el artículo 16 establece que, "…El tratamiento de los datos personales sensibles sólo puede realizarse cuando el titular a quien conciernen estos datos manifiesta su consentimiento en forma expresa, otorgado a través de una declaración escrita, verbal o por un medio tecnológico equivalente". Por su parte, el artículo 16. Ter. sostiene que los Datos personales biométricos

> …Son datos personales biométricos aquellos obtenidos a partir de un tratamiento técnico específico, relativos a las características físicas, fisiológicas o conductuales

de una persona que permitan o confirmen la identificación única de ella, tales como la huella digital, el iris, los rasgos de la mano o faciales y la voz" y agrega que, "Sólo podrán tratarse estos datos cuando se cumpliere con lo dispuesto en el inciso primero del artículo 16 y siempre que el responsable proporcione al titular la siguiente información específica: a) La identificación del sistema biométrico usado; b) La finalidad específica para la cual los datos recolectados por el sistema biométrico serán utilizados; c) El período durante el cual los datos biométricos serán utilizados, y d) La forma en que el titular puede ejercer sus derechos.

3.3. Definiciones

En lo relativo a las definiciones, el artículo 3 del Proyecto de Ley dispone:

Definiciones. Para los efectos de la presente ley, se entenderá por:

a) Imitación digital realista: reproducción audiovisual o sonora realista de una persona mediante IA, que simula su voz, rostro, cuerpo o expresión.

b) Contenido manipulado: contenido digital alterado para aparentar que una persona dijo o hizo algo que no ocurrió.

c) Consentimiento: autorización previa, expresa, informada y verificable para el uso de imagen, voz o cuerpo mediante IA. Es revocable en cualquier momento.

d) Contenido exento: representaciones claramente identificables como parodia, sátira, crítica o expresión artística, siempre que no causen daño grave o desinformación.

e) Plataformas digitales: cualquier servicio que aloje, indexe o difunda contenidos generados por terceros.

En su conjunto, las definiciones presentan una excesiva amplitud e indeterminación, con solapamientos internos y conceptos abiertos sin criterios de cierre, lo que debilita su coherencia y previsibilidad. Así, se describen fenómenos tecnológicos más que categorías jurídicas, sin atender suficientemente a su relevancia funcional para el daño o la imputación de responsabilidad, trasladando el peso regulatorio al intérprete y generando riesgos de inseguridad jurídica y sobrerregulación.

En cualquier caso, nos parece pertinente destacar que, la definición de "contenido exento" de la letra d) merece una consideración adicional desde la perspectiva de la libertad de expresión política. Al condicionar la exención a que el contenido sea "claramente identificable" como parodia o crítica, y a que no cause "daño grave", el proyecto traslada al intérprete —y eventualmente al juez o regulador— la decisión sobre si una imitación digital de un político constituye sátira legítima o infracción sancionable. La indeterminación de

ambos estándares puede producir un efecto de desaliento (*chilling effect*) sobre quienes crean contenido crítico o humorístico de figuras públicas. Así, ante la incerteza sobre si su obra quedará amparada por la exención, lo racional, pareciera, es abstenerse. Esta tensión no es teórica. La experiencia comparada muestra que leyes de regulación de *deepfakes* con cláusulas de excepción análogas han sido impugnadas y, en varios casos, invalidadas precisamente porque sus estándares vagos desincentivan la sátira y el comentario político, formas de expresión que históricamente han constituido herramientas fundamentales de fiscalización del poder[12].

4. El derecho a la integridad digital

4.1. Regulación en el proyecto de ley

El título II "Del derecho a la integridad digital", se regula en los artículos 4 a 8 del Proyecto de Ley.

El artículo 4. dispone: "Derecho exclusivo. Toda persona tiene el derecho exclusivo a autorizar la reproducción digital de su imagen, cuerpo o voz mediante IA. El uso no autorizado constituye una infracción a la integridad digital."

Por su parte, los artículos 5 a 8 configuran un régimen de prohibición y excepciones en torno a las imitaciones digitales realistas: el artículo 5 establece como regla general la prohibición de generar, difundir o almacenar este tipo de contenidos sin consentimiento; el artículo 6 introduce excepciones limitadas para fines satíricos, críticos o informativos, siempre que el contenido sea manifiestamente irreal y no produzca desinformación ni perjuicio grave; el artículo 7 extiende la protección de este derecho hasta cincuenta años después del fallecimiento de la persona imitada, atribuyendo el consentimiento a herederos o representantes; y el artículo 8 refuerza la tutela respecto de artistas y figuras públicas, exigiendo consentimiento para reproducciones realistas mediante IA cuando no concurra una finalidad informativa o crítica.

4.2. Que es el derecho a la integridad digital

Actualmente, se ha trabajado la idea de la integridad digital de la persona como fundamento de los derechos digitales, cuyos autores, vinculados a la Universidad Palacký de Olomouc, sostienen que el derecho fundamental a la protección de datos personales es insuficiente para proteger efectivamente a la persona en la era digital, pues funciona principalmente como un instrumento de regulación del mercado y de facilitación del libre flujo de datos, más que como un derecho centrado en la integridad personal (Vardanyan et al., 2022, pp. 163-167); esta orientación —agrega el informe— favorece la comodificación de los datos personales y deja sin tutela adecuada la dimensión identitaria del individuo, al

[12] Sobre la tensión entre la regulación de los *deepfakes* y la libertad de expresión: Broinowski y Martin, 2024, pp. 2575-2594.

proteger el dato como objeto autónomo y no como proyección de la persona (Vardanyan et al., 2022, pp. 168-171).

Frente a ello, el trabajo de dichos autores propone reconocer un derecho a la integridad digital como extensión de la integridad personal, capaz de abarcar la "vida digital" y la noción de un sujeto digital cuya identidad se construye a partir de fragmentos informacionales (Vardanyan et al., 2022, pp. 170-173).

Este nuevo derecho se fundamenta, según concluyen los autores, en una concepción restrictiva de la dignidad humana, entendida como límite a la autonomía y al mercado, destinada a impedir la instrumentalización y apropiación de la identidad personal en el entorno digital (Vardanyan et al., 2022, pp. 174-178), y permitiría —según los autores— reconfigurar la protección de los derechos digitales colocando nuevamente a la persona —y no al dato— en el centro del sistema jurídico (Vardanyan et al., 2022, pp. 179-180).Lo anterior, ha tenido manifestaciones políticas y jurídicas concretas. Así, el Consejo Municipal de Estrasburgo aprobó en diciembre de 2024 una moción que reconoce el derecho a la integridad digital y exige que la digitalización de los servicios públicos no excluya a las personas, garantizando siempre alternativas presenciales y acompañamiento humano, promoviendo la inclusión digital y evitando que el uso obligatorio de tecnologías vulnere el acceso efectivo a derechos y servicios públicos (Cf. Ville de Strasbourg, 2024). Asimismo, el 28 de septiembre de 2025, en una votación de democracia directa con una participación cercana al 50%, la ciudadanía suiza aprobó por estrecho margen la introducción de un sistema de identidad digital oficial, voluntario y gestionado por el Estado (50,4%), que coexistirá con el pasaporte y la cédula física y cuya implementación está prevista para 2026 (Cf. Stephens et al., 2025).

En este contexto, el derecho a la integridad digital importaría una novedad en nuestro ordenamiento jurídico, el que no se encuentra reconocido directamente por la Constitución Política de la República ni por ninguna ley general o especial. Esto tendría especial relevancia si sostenemos, siguiendo el lineamiento europeo, que la integridad digital opera como una categoría autónoma e independiente de los demás derechos como la integridad psíquica o la vida privada.

5. Responsabilidades y sanciones

El título III del Proyecto de Ley contempla las "Responsabilidades y sanciones" en tres artículos, que regulan las obligaciones de las plataformas digitales (artículo 9); las sanciones por infracción de la ley (artículo 10); y, las medidas cautelares para el retiro o bloqueo del contenido (artículo 11).

5.1. Obligaciones de plataformas digitales (artículo 9)

> Artículo 9.- Obligaciones de plataformas digitales. Las plataformas deberán retirar el contenido no autorizado en un plazo máximo de 72 horas desde la notificación

del afectado o sus representantes (inciso 1). La omisión será considerada infracción grave y sancionada administrativamente.

Las plataformas definidas, que, como se dijo, son definidas en el artículo 3, letra e) como: "cualquier servicio que aloje, indexe o difunda contenidos generados por terceros" tienen la obligación de retirar el contenido no autorizado en el plazo y bajo las sanciones que indica el artículo 9 precedente.

5.2. Sanciones por la infracción de la ley.

Artículo 10.- La infracción a esta ley será sancionada con:

a) Multa de 100 a 1.000 UTM,

b) Indemnización por daño moral y patrimonial,

c) Medidas de retiro forzoso del contenido, sin perjuicio de las acciones penales que correspondan por suplantación de identidad, injurias, calumnias o fraude.

Este artículo 10 presenta una grave imprecisión técnica y legislativa. En primer lugar, su epígrafe alude a las "sanciones" aplicables por infracción de la ley, cuando, en rigor, la única sanción propiamente tal es la prevista en la letra a), consistente en una multa de 100 a 1.000 UTM. En efecto, la indemnización de los daños moral y patrimonial no constituye una sanción, sino una medida de carácter reparatorio o resarcitorio, orientada a la compensación del daño efectivamente causado.

En este sentido, como sostiene San Martín, la pregunta central de la responsabilidad civil sobre por qué alguien debe indemnizar un daño solo se comprende a partir de las funciones que se le han atribuido históricamente, las cuales determinan su estructura y requisitos, destacando tres enfoques principales: la función punitiva o retributiva, centrada en la culpa del agente y propia de los sistemas subjetivos inspirados en los códigos decimonónicos; la función compensatoria o correctiva, que pone el acento en el daño injustamente sufrido por la víctima y explica el desarrollo de la responsabilidad objetiva basada en el riesgo; y la función preventiva o disuasiva, promovida por el análisis económico del derecho, que concibe la responsabilidad como un incentivo para evitar daños futuros asignando los costos al sujeto mejor situado para prevenirlos, subrayando el autor que la falta de claridad sobre qué función debe primar genera tensiones estructurales y prácticas, especialmente en sociedades influidas por el *principio pro damnato*, donde resulta incómoda la idea de que, actuando diligentemente, el daño pueda quedar a cargo de la víctima (Cf. San Martín, 2025).

La referida confusión, se ve reafirmada al analizar el artículo 11, que comprende las medidas de retiro forzoso del contenido mencionados en el artículo 10 letra C), siendo reiterativa e innecesaria su mención. Así, el artículo 11 dispone: "Medidas cautelares. El

tribunal podrá decretar, en cualquier momento, el retiro o bloqueo del contenido, a solicitud del afectado, mientras se resuelve el fondo del litigio".

Por lo demás, y en cuanto a la multa de 100 a 1000 UTM, el legislador no ha considerado, para establecer el monto de la multa, las capacidades económicas del responsable de la infracción, la relevancia social de la misma, o los lucros obtenidos o que razonablemente se pudieran haber incorporado por el infractor[13].

6. Conclusiones

1. La inteligencia artificial generativa no es una categoría jurídica autónoma, sino una manifestación de los modelos y sistemas de IA de uso general. El enfoque europeo, funcional y basado en el riesgo, distingue entre el modelo generativo y su despliegue como sistema, lo que permite asignar deberes y responsabilidades a lo largo de la cadena de valor, especialmente frente a riesgos como los *deepfakes*. La omisión de esta distinción en el proyecto de ley chileno constituye una debilidad estructural de su diseño normativo.

2. Los *deepfakes* constituyen una de las manifestaciones más disruptivas de la inteligencia artificial generativa, al permitir la creación de contenidos sintéticos altamente realistas con un elevado potencial de engaño y daño social, que afectan derechos como la privacidad, la dignidad, la reputación, la integridad informativa y el funcionamiento democrático. Mientras la Unión Europea ha respondido mediante un enfoque integrado y preventivo, basado en la transparencia y la gestión de riesgos, el ordenamiento chileno ha abordado el fenómeno de forma fragmentaria, con respuestas parciales y escasamente articuladas, lo que revela una evolución normativa aún incompleta y justifica la necesidad de un marco más coherente y funcional que permita distinguir usos legítimos y lesivos, asignar responsabilidades claras y asegurar una tutela efectiva frente a las imitaciones no consentidas generadas mediante inteligencia artificial.

3. El Boletín N°17.795-19 persigue un objetivo legítimo al proteger la identidad e integridad frente a imitaciones digitales no consentidas, pero presenta debilidades estructurales que afectan su eficacia. En particular, centra su objeto en el "uso" de la IA sin definirlo ni articularlo con la comercialización y la puesta en servicio; delimita el núcleo protegido excluyendo elementos relevantes del daño, lo que exige una compleja coordinación con el marco constitucional y la Ley 21.719; y adopta definiciones amplias e indeterminadas que describen fenómenos tecnológicos más que categorías jurídicas operativas. En conjunto, el proyecto supone un avance relevante, pero requiere mayor precisión conceptual y coherencia sistemática para cubrir el ciclo completo del riesgo y asegurar previsibilidad normativa.

[13] Estos mismos reproches ha observado la doctrina a propósito del artículo 24 de la Ley 19.496 (Munita Marambio, 2022, p. 612).

4. El Proyecto de Ley 17.795-19 introduce por primera vez en el derecho chileno el derecho a la integridad digital, configurándolo como un derecho exclusivo a autorizar la reproducción digital de la imagen, el cuerpo o la voz mediante IA, lo que supone una categoría autónoma de tutela centrada en la identidad personal. El régimen adoptado es predominantemente preventivo, al prohibir las imitaciones no consentidas, establecer excepciones restrictivas, extender la protección más allá de la muerte y reforzar la tutela de artistas y figuras públicas. No obstante, al no encontrarse este derecho expresamente reconocido en la Constitución ni en una ley marco general, su estatuto jurídico y su relación con derechos clásicos y con las libertades comunicativas quedan abiertos a interpretación, trasladando al intérprete la delimitación de un derecho aún en proceso de consolidación normativa.

5. El régimen de responsabilidades y sanciones del Proyecto de Ley presenta deficiencias estructurales y técnicas que afectan su coherencia y eficacia. En particular, el artículo 10 confunde sanción administrativa e indemnización de perjuicios, desdibujando la distinción entre función punitiva y responsabilidad civil resarcitoria, y reitera de forma imprecisa las medidas de retiro del contenido como sanción y como cautelar. A ello se suma un diseño de multas sin criterios de graduación vinculados a la gravedad de la conducta, la capacidad económica del infractor o los beneficios obtenidos. En conjunto, el título III refleja una regulación fragmentaria que requiere una revisión orientada a clarificar categorías, funciones y asegurar un sistema de sanciones y remedios coherente con los principios del derecho sancionador y de la responsabilidad civil.

Acerca del artículo

Notas de conflicto de interés. El autor declara no tener ningún conflicto de interés en relación con la publicación de este artículo.

Contribución en el trabajo. El autor asumió todos los roles establecidos en Contributor Roles Taxonomy (CRediT).

Bibliografía

Amunátegui Perelló, C. (2021). *Arcana Technicae. El Derecho y la Inteligencia Artificial.* Tirant lo Blanch.

Amunátegui Perelló, C. (2025, 4 de septiembre). Hacia un suicidio artificial. *El Mercurio*. https://www.elmercurio.com/legal/Registro/Login.aspx?urlBack=/Legal/Noticias/Opinion/2025/09/04/915346/hacia-un-suicidio-artificial.aspx

Aragüez Valenzuela, L. (2024). Hacia la eticidad algorítmica en las relaciones laborales. Ediciones Laborum.

Bentzen, N. (2025). *Information manipulation in the age of generative artificial intelligence.* European Parliamentary Research Service (EPRS). Members' Research Service.

Bertolini, A. (2025). ¡La inteligencia artificial no existe! Desafiando la narrativa de la neutralidad tecnológica en la regulación de la responsabilidad civil de las tecnologías avanzadas. *Revista de Derecho Privado, 49,* 31-82. https://doi.org/10.18601/01234366.49.02

Broinowski, A. y Martin, F. R. (2024). Beyond the deepfake problem: Benefits, risks and regulation of generative AI screen technologies. *Journalism Practice, SAGE Publications, 18*(10), 2575–2594.

Comisión Europea. (2018, 25 de abril). Comunicación de la Comisión al Parlamento Europeo, al Consejo Europeo, al Consejo, al Comité Económico y Social Europeo, y al Comité de las Regiones. Inteligencia Artificial para Europa.

Comisión Europea. (2020, 19 de febrero). *White Paper on Artificial Intelligence – A European approach to excellence and trust.* https://commission.europa.eu/publications/white-paper-artificial-intelligence-european-approach-excellence-and-trust_en

Comisión Europea, Joint Research Centre. (2025). *Generative AI outlook report 2025. Publications Office of the European Union*. https://doi.org/10.2760/104652

Dhiman, B. (2023). Exploding AI-Generated Deepfakes and Misinformation: A Threat of Global Concern in the 21st Century. *Qeios.* https://www.qeios.com/read/DPLE2L

Diez Ringele, F. (2024). Explorando la responsabilidad civil y el riesgo en el proyecto de ley chileno sobre inteligencia artificial*. Revista Actualidad Jurídica, 50.* https://doi.org/10.1177/1329878X241288034

European Data Protection Supervisor (2023). Deepfake detection. https://www.edps.europa.eu/data-protection/technology-monitoring/techsonar/deepfake-detection_en

Fernández, R. (2015, 20 de septiembre). El derecho a la propia imagen [Artículo de blog]. https://www.jraulfernandez.es/el-derecho-a-la-propia-imagen/

Fernández-Llorca, D., Gómez, E., Mazzini, G. y Sánchez, I. (2025). An interdisciplinary account of the terminological choices by EU policymakers ahead of the final agreement on the AI Act: AI system, general purpose AI system, foundation model, and generative AI. *Artificial Intelligence and Law, 33*, 875–888.

García de Blanes Sebastián, M., Díaz-Marcos, L., Aguado Tevar, Ó. y Delso Vicente, A. T. (2025). Análisis de las aplicaciones de la Inteligencia Artificial Generativa en sectores estratégicos: Una revisión de literatura. *Revista Latina de Comunicación Social, 83,* 1-24.

Jabbaz Rosenbaum, V. (2025). Deepfakes íntimos no consentidos: desafíos del ordenamiento jurídico chileno y el rol de la responsabilidad civil ante la IA. *Revista de Derecho Privado: Debates y Tendencias, 2*, 35-73.

Mahler, T. (2025). Risk Narrative: Deconstructing the AIA's Risk-Based Approach as a Regulatory Heuristic. En Raue, J., von Ungern-Sternberg, A., Kumkar, M. y Rüfner, T. (Eds.), *Artificial Intelligence and Fundamental Rights. The AI Act of the European Union and its implications for global technology regulation* (pp. 57-74). Institute for Digital Law Trier (IRDT).

Martín-Casals, M.(2022). Desarrollo tecnológico y responsabilidad extracontractual. A propósito de los sistemas de inteligencia artificial (IA). En J. Pérez Collados (Coord.), *La cultura jurídica en la era digital, Cuadernos Digitales. Derecho y Nuevas Tecnologías* (pp. 101-138). Thomson Reuters Aranzadi.

Masood, M., Nawaz, M., Malik, KM, Javed, A., Irtaza, A. y Malik, H. (2023). Generación y detección de deepfakes: Estado del arte, desafíos abiertos, contramedidas y futuro. *Inteligencia Aplicada, 53*(4), 3974-4026 .

Munita Marambio, R. (2022). Los daños punitivos y su tratamiento en la LPC. En E. Isler y F. Fernández (Dirs.), *GPS Consumo. Guía profesional.* Tirant-Lo Blanch

Navas Navarro, S. (2022). *Daños ocasionados por sistemas de inteligencia artificial. Especial atención a su futura regulación.* Editorial Comares.

Patil, K., Kale, S., Dhokey, J. y Gulhane, A. (2023). Detección de deepfakes mediante características biológicas: un estudio. Preimpresión de arXiv:2301.05819.

Ramos-Zaga, F. (2024). Deepfake: Análisis de sus implicancias tecnológicas y jurídicas en la era de la Inteligencia Artificial. *Derecho Global. Estudios sobre Derecho y Justicia*, 9(27), 359-387. https://doi.org/10.32870/dgedj.v9i27.754

Roberts, R. (2025). *Regulación extranjera sobre deepfakes. Estados Unidos, Francia, Australia, Dinamarca y otros* (Asesoría Técnica Parlamentaria, SUP N°151302). Biblioteca del Congreso Nacional de Chile. https://obtienearchivo.bcn.cl/obtienearchivo?id=repositorio/10221/37951/2/BCN_regulacion_deepfake_dic2025.pdf

Russell, S. J. y Norvig, P. (2022). *Artificial intelligence: A modern approach* (4.ª ed.). Pearson Education Limited.

San Martín, L. (2025, 25 de julio). Funciones y límites de la responsabilidad civil. *El Mercurio.* https://www.elmercurio.com/legal/analisis/index.aspx?userNameAutor=lsan%20mart%C3%ADn&tipoPortadilla=Autor

Stephens, T., Gysi von Wartburg, R. y Islas, P. (2025, 28 de septiembre). Suiza acepta un sistema de identidad digital y dice adiós a una tasa de uso de propiedad residencial. *SWI.* https://www.swissinfo.ch

Trinh, L. y Liu, Y. (2021). Un análisis de la imparcialidad de los modelos de IA para la detección de deepfakes. Preimpresión de arXiv:2105.00558.

Vardanyan, L., Stehlík, V, Kocharyan, H. (2022). Digital Integrity: A Foundation for Digital Rights and the New Manifestation of Human Dignity. *TalTech Journal of European Studies, 12*(1), 159-180. https://doi.org/10.2478/bjes-2022-0008

Ville de Strasbourg (2024, 9 de diciembre). Droit à l'intégrité numérique et préservation de l'accès aux services publics. https://www.strasbourg.eu

Von Ungern-Sternberg, A. (2025). Freedom of Expression and the Regulation of AI. En Raue, J., von Ungern-Sternberg, A., Kumkar, M. y Rüfner, T. (Eds.), *Artificial Intelligence and Fundamental Rights. The AI Act of the European Union and its Implications for Global Technology Regulation* (pp. 73-93). Institute for Digital Law Trier (IRDT).

Normativa

Chilena

Cámara de Diputadas y Diputados de Chile. Proyecto de ley que regula la creación y difusión de imitaciones digitales realistas de la imagen, cuerpo o voz de las personas, generadas mediante inteligencia artificial, Boletín Nº17.795-19, ingresado el 21 de agosto de 2025, Primer trámite constitucional, Cámara de Diputados.

Constitución Política de la República de Chile.

Código Penal. (12 de noviembre de 1874). https://bcn.cl/2f6m7

Ley 21.719. (13 de diciembre de 2024). Que regula la protección y el tratamiento de los datos personales. https://bcn.cl/GapReB

Senado de la República de Chile. Proyecto de ley que regula los sistemas de inteligencia artificial, Boletín Nº16.281-19, ingresado el 7 de mayo de 2024, Primer trámite constitucional.

Senado de la República de Chile. Proyecto de ley que modifica el Código Penal, con el objeto de tipificar la generación y difusión de imágenes o hechos de carácter privado o íntimo, creados con herramientas de inteligencia artificial, Boletín Nº17.307-07, ingresado el 17 de diciembre de 2024, Primer trámite constitucional.

Extranjera

Comisión Europea. (2025, 6 de febrero). Comunicación a la Comisión. Aprobación del contenido del proyecto de Comunicación de la Comisión - Directrices de la Comisión sobre la definición de sistema de inteligencia artificial establecidas por el Reglamento (UE) 2024/1689 [Ley de IA]. https://artificialintelligenceact.eu/es/ai-act-explorer/

Reglamento (UE) 2022/2065 del Parlamento Europeo y del Consejo, de 19 de octubre de 2022, relativo a un mercado único de servicios digitales y por el que se modifica la Directiva 2000/31/CE [Reglamento de Servicios Digitales – DSA].

Reglamento (UE) 2024/1689 del Parlamento Europeo y del Consejo de 13 de junio de 2024 por el que se establecen normas armonizadas en materia de inteligencia artificial. https://eur-lex.europa.eu/legal-content/ES/TXT/PDF/?uri=OJ:L_202401689.

Tools to Address Known Exploitation by Immobilizing Technological Deepfakes on Websites and Networks Act [*Take It Down Act*], Pub. L. No. 119-12, 139 Stat. 55 (2025). https://c.bcn.cl/rxhgg3