UNIVERSIDAD FINIS TERRAE
FACULTAD DE DERECHO
ESCUELA DE DERECHO

RESPONSABILIDAD CIVIL POR DAÑOS DERIVADOS DEL USO DE INTELIGENCIA ARTIFICIAL

REINALDO ESCOBAR TORREJÓN

Memoria presentada a la Facultad de Derecho de la Universidad Finis Terrae, para optar al grado de Licenciado en Ciencias Jurídicas

Profesora Guía: Isabel Wigg Sotomayor

Santiago, Chile
2026

DEDICATORIA

Gracias…

M.R.T.B, por enseñarme la pasión;

M.C.G.G., por entregarme lealtad;

C.A.T.G., por tu incondicionalidad;

V.C.T.G., por demostrarme lo que es la valentía y fortaleza;

I.W.S., por creer en mí;

J.Z.H., por todo...

ÍNDICE

INTRODUCCIÓN **6**

CAPÍTULO 1: EL AUGE DE LA INTELIGENCIA ARTIFICIAL EN EL DERECHO **9**

1.1. Origen y evolución histórica de la IA 9

1.2. Concepto de la IA 12

1.3. Experiencias internacionales en la regulación de la IA 14

1.4. Conclusiones preliminares 26

CAPÍTULO 2: EL DERECHO DE DAÑOS ANTE LA INTELIGENCIA ARTIFICIAL **29**

2.1. La función preventiva de la responsabilidad civil en el contexto tecnológico 29

2.2. El principio de la reparación integral de los daños derivados en uso de IA 31

2.3. Aplicación de los sistemas de atribución de responsabilidad civil 34

2.4. Régimen de responsabilidad civil en el proyecto de ley que regula la IA en Chile 42

2.5. Conclusiones preliminares 47

CAPÍTULO 3: IA Y RESPONSABILIDAD EN LA LABOR JURÍDICA **49**

3.1. Los sistemas expertos o sistemas basados en conocimiento 49

3.2. Responsabilidad de los abogados por los daños derivados del uso negligente de IA en el ejercicio de la profesión 51

3.3. Utilización de la IA en el ejercicio de la jurisdicción 59

3.4. Conclusiones preliminares 69

CONCLUSIÓN **71**

BIBLIOGRAFÍA **75**

RESUMEN

La presente investigación analiza el problema de la atribución de responsabilidad civil por los daños derivados del uso de inteligencia artificial en contextos jurídicamente relevantes, con el objetivo de examinar si las categorías clásicas del derecho de daños ofrecen herramientas suficientes frente a procesos algorítmicos opacos. Para ello, se emplea un análisis que combina el estudio de fuentes normativas y doctrinarias con la revisión de experiencias regulatorias comparadas, examinando las clasificaciones existentes sobre sistemas de IA, los distintos regímenes de atribución de responsabilidad civil y el proyecto de ley que regula la IA en Chile. Asimismo, se aborda la proyección de estas cuestiones sobre el ejercicio profesional de abogados y jueces, considerando los deberes de diligencia que rigen su labor al utilizar estas herramientas. El trabajo discute las distintas posturas doctrinarias en torno a los modelos de atribución de responsabilidad y al tratamiento de la opacidad algorítmica frente a la reconstrucción del nexo causal. A partir de este análisis, se plantean criterios para orientar la imputación de responsabilidad hacia los agentes humanos que intervienen en el diseño, desarrollo o uso de los sistemas de IA, contribuyendo a la discusión sobre la adecuación de las instituciones tradicionales de la responsabilidad civil frente a los desafíos que plantea esta tecnología.

Palabras Clave: Derecho de Daños, Error Judicial, Inteligencia Artificial, Lex Artis, Responsabilidad Civil Profesional.

ABSTRACT

This research analyzes the issue of attributing civil liability for damages arising from the use of artificial intelligence in legally relevant contexts, with the aim of examining whether the traditional categories of tort law offer sufficient tools for addressing opaque algorithmic processes. To this end, the study employs an analysis that combines the examination of normative and doctrinal sources with a review of comparative regulatory experiences, examining existing classifications of AI systems, the various regimes for attributing civil liability, and the draft bill regulating AI in Chile. It also addresses the implications of these issues for the professional practice of lawyers and judges, considering the duties of diligence that govern their work when using these tools. The study discusses the various doctrinal positions regarding models of liability attribution and the treatment of algorithmic opacity in relation to the reconstruction of the causal link. Based on this analysis, criteria are proposed to guide the attribution of liability to human agents involved in the design, development, or use of AI systems, thereby contributing to the discussion on the adequacy of traditional civil liability institutions in the face of the challenges posed by this technology.

Keywords: Tort Law, Judicial Error, Artificial Intelligence, Lex Artis, Professional Liability.

INTRODUCCIÓN

El derecho no es estático. Nunca lo ha sido. Cada vez que la sociedad se transforma, el ordenamiento jurídico se ve obligado a reexaminar sus categorías, revisar sus fundamentos y encontrar respuestas para problemas que sus creadores no anticiparon. La introducción masiva de la Inteligencia Artificial (en adelante, IA) en los diferentes ámbitos de nuestras vidas, representa, a mi juicio, uno de los desafíos más importantes que el derecho contemporáneo ha debido enfrentar, ya que, su velocidad de desarrollo y su vinculación en ámbitos sensibles ha superado, con creces, la capacidad de respuesta de los ordenamientos tradicionales.

En este sentido, el origen de esta investigación parte de un interés genuinamente personal pero evidentemente social. Me preocupa que la complejidad tecnológica pueda convertirse en un escudo detrás del cual la responsabilidad humana se diluya hasta desaparecer. La pregunta que organiza este trabajo es: cuando el uso de IA en el ámbito jurídico produce un perjuicio, ¿quién debe responder? Detrás de esa interrogante hay una cuestión más profunda, que tiene que ver con los valores que el derecho de daños ha protegido históricamente; la dignidad de quien resulta perjudicado, la exigencia de que quien causa un daño lo repare, y la convicción de que la complejidad de un instrumento no puede ser argumento para eludir las consecuencias de haberlo usado mal.

El problema central que esta investigación aborda no es, por tanto, la IA en sí misma, sino la responsabilidad civil derivada de su uso en contextos jurídicamente relevantes. En particular, me interesa examinar si las categorías tradicionales del derecho de daños, construidas sobre la base de conductas humanas identificables y nexos causales reconstruibles, ofrecen suficientes herramientas para enfrentar escenarios en los que la decisión que produjo el perjuicio ocurrió, en todo o en parte, dentro de un proceso algorítmico opaco. Esta pregunta adquiere especial urgencia cuando la IA comienza a intervenir dentro del sistema de justicia, en que los bienes jurídicos comprometidos incluyen, más que intereses patrimoniales, derechos fundamentales cuya protección el ordenamiento no puede relativizar.

La postura que sostengo parte de una premisa que considero irrenunciable, que la IA no puede ser sujeto de responsabilidad jurídica. Carece de conciencia, de voluntad moral y de capacidad de autodeterminación, por lo que atribuirle personalidad jurídica no resolvería el problema de la imputación, sino que lo relegaría, creando un centro ficticio de responsabilidad que, en la práctica, podría operar como pantalla para quienes verdaderamente deben responder. Por eso, mi argumento central es que toda acción dañosa vinculada al uso de IA debe ser reconducida, necesariamente, hacia los humanos que intervinieron en su diseño, desarrollo, implementación o uso. El foco del análisis no puede estar en la máquina; debe estar en las personas que decidieron usarla y en el estándar de diligencia que se les exigía.

Este trabajo exigió, desde el inicio, un análisis jurídico dogmático, atento también al fenómeno tecnológico que lo origina. La responsabilidad civil frente a la IA es, sobre todo, una construcción conceptual. Pero no puede prescindir de ciertos elementos empíricos, tales como el desarrollo técnico de estos sistemas, las respuestas regulatorias adoptadas en distintos países y la jurisprudencia que recién comienza a formarse. Por eso fue necesario analizar el origen, precisar su naturaleza jurídica y ordenar los distintos modelos regulatorios que el derecho comparado ha propuesto, antes de proyectar todo ese análisis sobre el sistema de responsabilidad civil.

La investigación se basó en fuentes normativas, doctrinales y jurisprudenciales, haciendo de ellas una lectura reflexiva, para extraer el núcleo dogmático de la responsabilidad civil aplicada a la IA, precisar sus presupuestos de procedencia y establecer criterios razonables de atribución. Busco identificar algunos de los problemas que surgen al aplicar las reglas tradicionales de responsabilidad a un fenómeno nuevo, rescatar los aportes de cada modelo regulatorio y ordenar un marco que dé coherencia a la solución de estos casos.

Como se trata de una materia con desarrollo desigual entre países, recurrí al método comparado, revisando las experiencias, bastante diversas, de la Unión Europea y Estados Unidos, para luego analizar el estado actual del proceso regulatorio en Chile. Esta comparación no buscó trazar paralelismos formales, sino identificar los problemas comunes a distintos sistemas y ver cómo cada uno ha intentado resolverlos. Sirvió, en definitiva, para mostrar los desafíos estructurales que plantea incorporar la IA al derecho de daños, más allá de cuánto haya avanzado la normativa en cada país.

La investigación siguió un orden progresivo. Primero fue necesario fijar el marco general: su origen, su naturaleza jurídica y los principales modelos regulatorios. Luego, el análisis se centró en el derecho de daños propiamente tal, aplicando la vigencia de las reglas tradicionales de responsabilidad civil, las funciones que este sistema debe cumplir frente a la IA y los regímenes de atribución disponibles. Finalmente, reflexioné sobre el uso de la IA en la administración de justicia, considerando las dificultades que plantea para el ejercicio profesional y la función jurisdiccional.

A lo largo de todo el trabajo, busqué mantener una línea de razonamiento clara, en que cada parte del análisis preparara la siguiente. El esfuerzo metodológico consistió, en ese sentido, en ordenar los problemas, exponer sus puntos críticos y construir un esquema que permita comprender y resolver con mayor seguridad los conflictos de responsabilidad civil vinculados a la IA.

Parto, en definitiva, de la convicción de que el derecho no puede quedar como observador pasivo frente a la transformación que la IA está produciendo. Su función no se agota en reaccionar ante el daño ya ocurrido; exige también anticipar riesgos, distribuir responsabilidades con equidad y proteger el lugar central de la persona humana en los procesos que el derecho regula. En el fondo, la discusión sobre IA y responsabilidad no trata de tecnología, sino de quién responde, ante quién y por qué. Y esa pregunta, como todas las que de verdad importan en el derecho, tiene siempre un rostro humano.

CAPÍTULO 1: EL AUGE DE LA INTELIGENCIA ARTIFICIAL EN EL DERECHO

1.1. Origen y evolución histórica de la IA

El origen de la IA se sitúa en la antigua Grecia, en que Aristóteles crea y establece los silogismos como un método de razonamiento deductivo e inductivo (Quintana, 2011, pp. 304-309). El mecanismo de los silogismos, fue recogido de forma posterior por la computación para el desarrollo de la IA bajo el nombre de algoritmos (Mancilla et al., 2016, pp. 1-8). Desde la segunda mitad del siglo XX, el desarrollo de estos algoritmos comenzó a adquirir una relevancia creciente dentro del ámbito científico y tecnológico, convirtiéndose en un elemento fundamental para el avance de la computación moderna (Rodríguez, 2018, pp. 25-47).

Uno de los antecedentes más importantes de su aplicación práctica se sitúa en la Segunda Guerra Mundial, particularmente en el uso del sistema de cifrado empleado por las fuerzas militares alemanas conocido como *Enigma*. Este sistema funcionaba mediante complejos procedimientos algorítmicos destinados a codificar y transmitir mensajes secretos entre las unidades militares. Su importancia histórica es que su sistema de cifrado fue finalmente descifrado gracias al trabajo de un grupo de matemáticos británicos dirigidos por Alan Turing, quienes desarrollaron métodos algorítmicos y dispositivos computacionales capaces de vulnerar dicho sistema criptográfico (Leavitt, 2006, pp. 181-216). Este logro no solo contribuyó al esfuerzo bélico aliado, sino que también sentó las bases fundamentales para el posterior desarrollo de la informática y de la teoría de la computación (Amunategui, 2020, pp. 30-32).

Estos avances condujeron a la consolidación de dos grandes corrientes dentro del desarrollo de la IA. Por un lado, la escuela simbolista, centrada en la representación formal del conocimiento y en el uso de reglas lógicas para el procesamiento de información; y, por otro lado, la escuela conexionista, basada en modelos inspirados en el funcionamiento de las redes neuronales y en los mecanismos de aprendizaje a partir de datos (Ballesteros, 1992, pp. 343-351). Estas corrientes han

influido en la evolución de la IA contemporánea y continúan orientando gran parte de la investigación en este ámbito.

Antes de que estas terminaran de consolidarse, ya existía un antecedente decisivo que había puesto sobre la mesa la pregunta de fondo que ambas escuelas intentarían responder desde enfoques distintos: si una máquina podía, de algún modo, pensar. En 1950, Alan Turing publicó el influyente artículo titulado *Computing Machinery and Intelligence*, texto que marcó un punto de inflexión en la reflexión académica sobre la relación entre las máquinas y la inteligencia humana. En este trabajo, el autor analizó la posibilidad de que los sistemas computacionales pudieran desarrollar comportamientos comparables a los del razonamiento humano, particularmente en el ámbito del procesamiento de información y la toma de decisiones (Allen, 1997, pp. 254–275).

A partir de lo anterior, se definieron diversas líneas de investigación dentro del campo de la IA. Entre estas destacó el estudio de las redes neuronales, concebidas como herramientas destinadas a la automatización de procesos complejos y a la resolución de problemas que hasta entonces, habían sido considerados exclusivamente propios de la capacidad intelectual humana (Navas, 2017, pp. 11 y ss.). Con ello, se permitió ampliar las posibilidades de desarrollo tecnológico y abrió nuevas perspectivas en la investigación sobre sistemas inteligentes.

En lo que se refiere al uso de la IA en el ámbito jurídico, se postularon diversos modelos computacionales que se pudiesen aplicar al razonamiento jurídico conforme a los principios de la escuela simbolista (Ashley, 2017, pp. 210-212). Sin embargo, estas iniciativas no alcanzaron los resultados esperados, ya que se trataba de diseños poco prácticos y altamente costosos. La principal dificultad radicó en la complejidad de traducir los enunciados normativos en algoritmos computacionales, debido a la ambigüedad y flexibilidad con que suelen redactarse las normas jurídicas (Amunategui, 2020, pp. 13-18).

El fracaso de estas primeras experiencias es el reflejo de un problema que terminó por frenar el avance de la IA aplicada al derecho. Este escenario contribuyó al surgimiento de un período conocido como el "invierno de la IA", que en el ámbito jurídico se tradujo en una notoria reducción del financiamiento y el interés por la investigación en esta materia (Valbuena, 2021, p. 33). Durante esta etapa se hizo evidente que la escuela simbolista presentaba importantes limitaciones para el desarrollo efectivo de sistemas de IA orientados al razonamiento legal (Belda, 2019, pp. 80-85),

permitiendo que la escuela conectivista adquiriera mayor relevancia en el desarrollo de estas tecnologías, incluidas aquellas destinadas a su aplicación en el campo jurídico.

En este contexto, el profesor David Rumelhart publicó un trabajo científico inspirado en las ideas previamente formuladas por Frank Rosenblatt. Se propuso perfeccionar el modelo original de redes neuronales mediante la incorporación de tres niveles de neuronas: una de entrada, una o varias capas ocultas o intermedias y una capa de salida, todas conectadas entre sí (Caicedo & López, 2009, p. 76). En otras palabras, el conjunto de datos que entran y salen de un sistema informático se transforma por etapas. Esta estructura permitiría que el sistema analizara la información ingresada de forma progresiva a través de diferentes filtros, lo que posibilitaría que el resultado final del algoritmo fuese producto del procesamiento realizado en cada uno de estos niveles de análisis.

Lo interesante, a mi juicio, está en cómo se entrena esa estructura. Antes de Rumelhart, el problema era que nadie sabía calcular el error de la capa oculta, porque no existía un valor esperado conocido para ella (Caicedo & López, 2009, pp. 74-75). Sin embargo, propone un algoritmo de retropropagación; primero se calcula el error en la salida, que sí es conocida, y luego ese error se devuelve hacia atrás, capa por capa, para estimar cuánto aportó cada neurona oculta al resultado y ajustar sus pesos (Caicedo & López, 2009, p. 77). En otras palabras, cada capa queda matemáticamente asociada a una porción del error. No es una operación única e indivisible, sino una secuencia de pasos que, en teoría, pueden rastrearse uno por uno.

Ahí veo la utilidad concreta para mi postura. Si el resultado de un sistema de IA es producto de etapas identificables y no de una caja cerrada, en principio es posible ubicar en qué capa se originó o se amplificó el error que terminó causando el daño. Eso le da un asidero técnico a la causalidad y el resultado dañoso deja de ser una incógnita total y pasa a ser, al menos parcialmente, un proceso descomponible. Lo mismo aplica para la atribución de responsabilidad. Si el defecto puede ubicarse en una etapa concreta del procesamiento, es más fácil determinar si el problema vino del diseño de la arquitectura, de los datos de entrenamiento o del usuario, y con ello, a cuál de los sujetos de la cadena corresponde responder.

Finalmente, y antes de conceptualizar la IA y cuáles son sus principales elementos, podemos concluir que un dispositivo puede considerarse inteligente cuando posee en su interior un

conjunto de algoritmos que le permiten ejecutar procesos de aprendizaje y razonamiento (Suñé Llinás, 2020, pp. 42-44). Incluso, hay quienes han advertido que estos sistemas podrían llegar a experimentar sensaciones y actuar de manera similar a los seres humanos (Manheim & Kaplan, 2019, pp. 111-113).

1.2. Concepto de la IA

La Real Academia Española define la IA como una "disciplina científica que se ocupa de crear programas informáticos que ejecutan operaciones comparables a las que realiza la mente humana, como el aprendizaje o el razonamiento lógico". Este concepto resalta principalmente el carácter científico y funcional, subrayando que su desarrollo no se limita a la simple automatización de tareas mecánicas. Por el contrario, busca replicar procesos cognitivos complejos que tradicionalmente han sido asociados con la actividad intelectual humana, tales como la capacidad de inferir, tomar decisiones o aprender a partir de la información disponible. Por otra parte, el Reglamento (UE) 2024/1689, conocido como la Ley de Inteligencia Artificial de la Unión Europea (en adelante, Reglamento UE), establece una definición operativa del concepto de sistema de IA dentro del ámbito regulatorio. En este sentido, el artículo 3º, numeral 1, señala:

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.

Con todo, la formulación de un concepto universal de IA presenta importantes dificultades. El acelerado desarrollo tecnológico y la complejidad para identificar rasgos esenciales y permanentes de esta tecnología provocan que muchas definiciones propuestas queden rápidamente superadas por la evolución de estos propios sistemas (Atienza, 2022, pp. 67-68).

En este sentido, parte de la doctrina ha manifestado una postura crítica frente a la pretensión de construir una definición unitaria. Así, autores sostienen que la heterogeneidad funcional y técnica de estas tecnologías impide establecer un concepto cerrado y definitivo. En consecuencia, para efectos del presente análisis tomaré como referencia la noción previamente expuesta, que se estructura en torno a la idea central de sistemas capaces de reproducir o emular determinados comportamientos o actividades que tradicionalmente se asocian a la inteligencia humana.

Por otra parte, conviene distinguir dos modalidades fundamentales. La doctrina suele diferenciar entre "IA débil" e "IA fuerte". De acuerdo con lo expuesto por Atienza (2022, pp. 67-68), la primera corresponde a aquellos sistemas capaces de adoptar decisiones de forma aparentemente autónoma, porque ha sido programado previamente para ello, con base en algoritmos que procesan grandes volúmenes de información, comúnmente denominados *big data*, cuya magnitud excede ampliamente la capacidad de análisis humano. Lo determinante es que la decisión adoptada no se debe a un proceso de aprendizaje, sino a la programación del robot a partir de una base de datos preexistente.

En contraste, la IA fuerte se refiere a sistemas cuyo nivel de sofisticación permite desarrollar procesos decisionales con un grado mucho mayor de autonomía. En estos casos, el algoritmo no se limita a ejecutar instrucciones predeterminadas, sino que es capaz de aprender a partir de la interacción con su entorno y de la continua recopilación de información, pudiendo adoptar decisiones o ejecutar acciones sin una intervención directa de sus programadores.

Finalmente, Ataz (2020, pp. 69-70), advierte que estos sistemas pueden operar con tal autonomía que ni sus propios desarrolladores lograrían explicar con exactitud los motivos de sus decisiones. Esta característica se vincula directamente con el problema de la "caja negra", expresión utilizada para describir la opacidad que puede presentar el funcionamiento interno de determinados sistemas algorítmicos. En tales circunstancias, la falta de transparencia en los procesos de decisión del sistema experto dificulta considerablemente la identificación de las causas que originaron un eventual daño, lo que plantea importantes desafíos para los regímenes tradicionales de atribución de responsabilidad (Scherer, 2016, pp. 353–400).

1.3. Experiencias internacionales en la regulación de la IA

1.3.1. Unión Europea

La regulación de la UE, se materializó en el Reglamento 2024/1689 (Parlamento Europeo y Consejo de la Unión Europea, 2024). Esta normativa representa uno de los primeros marcos jurídicos integrales a nivel mundial destinado a regular el desarrollo y uso de estas tecnologías. Su objetivo primordial es armonizar el funcionamiento del mercado interior mientras se garantiza un elevado nivel de protección de la salud, la seguridad y los derechos fundamentales frente a los efectos perjudiciales que los sistemas de IA pueden generar. Esta regulación se basa en el riesgo, clasificando las aplicaciones de IA en diferentes categorías, inaceptable, alto, limitado y mínimo, lo que determina la intensidad de las obligaciones y la responsabilidad exigible a los operadores.

Dentro de esta estructura, los sistemas de riesgo inaceptable son aquellos cuyas prácticas están estrictamente prohibidas por atentar contra los valores de la Unión, tales como la dignidad humana y la democracia. Entre estas conductas se encuentran la manipulación cognitiva del comportamiento de personas vulnerables, la clasificación social y el uso de sistemas de identificación biométrica remota en tiempo real en espacios públicos, salvo excepciones legales específicas. En este contexto, el daño se vincula a la pérdida de la capacidad de tomar decisiones informadas o a la vulneración de los derechos fundamentales, estableciendo que tales prácticas son incompatibles con el ordenamiento jurídico europeo.

Por otro lado, aquellos de alto riesgo constituyen el núcleo más detallado de la regulación, abarcando áreas como la infraestructura digital, la educación, el empleo, el acceso a servicios públicos y la aplicación de la ley. En estos casos, la responsabilidad de los proveedores es exhaustiva, exigiéndoles la implementación de un mecanismo de control y prevención de daños potenciales durante todo el ciclo de vida del producto. Como señalan Campusano y Giesen (2024, pp. 105-136), las obligaciones incluyen una gobernanza de datos rigurosa, documentación técnica,

registro de eventos y la garantía de una supervisión humana efectiva para prevenir resultados discriminatorios o lesivos.

El Reglamento UE configura un régimen de responsabilidad que, si bien se apoya en el marco administrativo y preventivo, interactúa directamente con el Derecho civil clásico. Por tanto, no se busca crear un nuevo estatuto de responsabilidad civil, sino complementar el Derecho vigente en materia de responsabilidad por daños causados por productos defectuosos, asegurando que los afectados por sistemas de IA dispongan de vías efectivas de reparación.

En este sentido, el considerando 10º del Reglamento aclara que los derechos de los consumidores y demás personas afectadas permanecen inalterados, permitiendo la reparación de potenciales perjuicios de conformidad con la Directiva 85/374/CEE del Consejo, de 25 de julio de 1985, relativa a la aproximación de las disposiciones legales, reglamentarias y administrativas de los Estados Miembros en materia de responsabilidad por los daños causados por productos defectuosos.

Dentro de la estructura de imputación del Reglamento, la figura del proveedor adquiere una especial relevancia para los sistemas de alto riesgo. El texto es enfático al señalar la necesidad de que una entidad asuma la carga de seguridad desde la fase de diseño. Al respecto, el Reglamento establece que:

> Conviene que una persona física o jurídica concreta, definida como el proveedor, asuma la responsabilidad asociada a la introducción en el mercado o la puesta en servicio de un sistema de IA de alto riesgo, con independencia de si dicha persona física o jurídica es o no quien diseñó o desarrolló el sistema (Parlamento Europeo y Consejo de la Unión Europea, 2024, p. 23, considerando Nº 79).

Esta disposición consagra una suerte de responsabilidad por la comercialización y puesta en marcha, vinculando la obligación de cumplir con los requisitos de seguridad a quien introduce la tecnología en el tráfico jurídico. Esto implica que el proveedor no puede exonerarse alegando que el algoritmo fue desarrollado por un tercero si él es responsable de su distribución en el mercado.

En definitiva, el Reglamento europeo, no regula la IA como un fenómeno estático, sino que distingue distintas etapas dentro de su ciclo de vida, atribuyendo a cada una exigencias de diligencia diferenciadas. En una primera fase, la norma reconoce un espacio de experimentación destinado a la investigación, el desarrollo y la creación de prototipos, incluyendo los denominados espacios controlados de pruebas para la IA y las pruebas en condiciones reales, en los que el sistema puede ser entrenado y validado antes de su introducción en el mercado (Reglamento UE 2024/1689, p. 92). En esta etapa, la exigencia de diligencia recae principalmente en el proveedor, quien debe adoptar desde el diseño un enfoque preventivo que anticipe los riesgos del sistema, implementando un sistema de gestión de riesgos que opere durante todo su ciclo de vida (Reglamento UE 2024/1689, p. 56).

Una segunda fase corresponde a la introducción en el mercado y la puesta en servicio propiamente tal, momento en que el sistema debe haber superado la correspondiente evaluación de conformidad a las obligaciones que el Reglamento impone a los proveedores (Reglamento UE 2024/1689, p. 62). En esta etapa se activa también la responsabilidad de otros actores de la cadena de valor, como los importadores, quienes deben verificar que el producto cuente con la documentación, el marcado y los requisitos exigidos antes de ponerlo a disposición en el mercado (Reglamento UE 2024/1689, p. 65).

Finalmente, una tercera fase se abre con la utilización efectiva del sistema, en la que el responsable del despliegue asume un deber de diligencia propio, vinculado al uso conforme a las instrucciones, la supervisión humana y la vigilancia continua de su funcionamiento (Reglamento UE 2024/1689, p. 67). Esta distribución progresiva de responsabilidades me permite sostener que el Reglamento no concentra el deber de cuidado en un único agente, sino que lo reparte a lo largo de toda la cadena, en función del momento del ciclo de vida en que cada actor interviene y del grado de control que efectivamente ejerce sobre el sistema.

Respecto al régimen sancionatorio, la normativa impone multas administrativas severas para desincentivar el incumplimiento. Las infracciones por prácticas prohibidas pueden alcanzar hasta los 35 millones de euros o el 7 % del volumen de negocios mundial anual de la empresa infractora, optándose por la cuantía más elevada. La magnitud de esta sanción refleja la gravedad con la que el Derecho europeo juzga la vulneración de los derechos fundamentales a través de la

IA, estableciendo un estándar de diligencia debida extremadamente alto para los operadores económicos, que podría tener como consecuencia un desincentivo al desarrollo.

Asimismo, el Reglamento define el daño en términos de "perjuicio considerable", el cual debe manifestarse en efectos adversos importantes en la salud física o mental, o en los intereses financieros de las personas. No obstante, según reporta Bermejo i Alegret, (2026), se ha debatido sobre el umbral de este daño, ya que el texto aprobado elevó la exigencia a perjuicios "considerables", lo que podría generar ciertas lagunas legales en comparación con propuestas iniciales menos restrictivas.

La Ley de IA de la Unión Europea no solo busca ordenar su mercado interno, sino consolidar el denominado "Efecto Bruselas", aspirando a convertirse en el estándar global de regulación ética. Su importancia trasciende sus fronteras al servir como modelo normativo para la mayoría de los países del centro y sur de América, en que se observan intentos de replicar esta estructura en proyectos legislativos recientes. La lección fundamental que esta regulación entrega a la comunidad internacional es la necesidad de transitar desde una ética puramente declarativa hacia una gobernanza institucional vinculante que, si bien promueve la innovación, establezca límites claros y responsabilidades exigibles para asegurar que el progreso tecnológico no se realice a expensas de las garantías constitucionales y la seguridad de los ciudadanos.

El proceso de elaboración del Reglamento UE se remonta al año 2015, cuando la Comisión Europea inicia los primeros trabajos preparatorios en un contexto tecnológico radicalmente distinto al que existiría al momento de su promulgación definitiva nueve años después. En aquella época, los sistemas de IA disponibles respondían a lo que técnicamente se conoce como IA estrecha o *narrow AI*: herramientas altamente especializadas, diseñadas para ejecutar una función acotada y predefinida con cierto grado de eficiencia, pero absolutamente incapaces de traspasar ese límite funcional (Nancholas, 2023, pp. 13-21). Un sistema de reconocimiento de imágenes, tal como su nombre lo sugería, reconocía imágenes, y nada más. Un algoritmo de recomendación musical solo sugería canciones, y así. La especialización era, estructuralmente, su condición de existencia.

El problema es que esa misma especificidad funcional fue la que le otorgó coherencia interna al esquema de clasificación de riesgos que el Reglamento adoptaría. Si un sistema solo puede hacer *una* cosa, resulta relativamente sencillo preguntarse si esa cosa es peligrosa, trivial o

intermedia, y regular en consecuencia. Sobre esa lógica se construyó la arquitectura normativa del reglamento; un sistema de categorías estratificadas que va desde la IA de riesgo mínimo, como un *chatbot* que recomienda canciones o una aplicación de filtros en fotografías, pasando por la IA de alto riesgo, que incluye sistemas que inciden en decisiones de salud, crédito, o administración de justicia, hasta llegar a las aplicaciones de riesgo inaceptable, cuyo uso queda prohibido, como los sistemas de manipulación conductual subliminal o el reconocimiento biométrico en espacios públicos con fines de vigilancia masiva. La clasificación es ordenada, sistemática y, en principio, razonable.

El inconveniente surge cuando se constata que el reglamento fue promulgado en 2024 y entró en vigencia progresivamente desde entonces en un ecosistema tecnológico que ya no se parece en nada al de 2015. La irrupción y masificación de los grandes modelos de lenguaje, como GPT-4 o sus sucesores, introdujo en el mercado una categoría de herramientas que el esquema regulatorio simplemente no anticipó. Estos sistemas no hacen una cosa, pueden hacer casi cualquier cosa, dependiendo de cómo se les instruya. El mismo modelo que hoy le sugiere una canción al usuario, asiste a un juez en la redacción de una sentencia, o puede ser integrado en un sistema de reconocimiento facial. La función no está inscrita en la herramienta, sino que depende del *prompt.* Y eso hace que la pregunta sobre qué nivel de riesgo representa este sistema, se vuelva, en la práctica, irresoluble bajo los parámetros del propio reglamento.

La consecuencia jurídica de este desajuste es de una gravedad que no debería subestimarse. ¿Bajo qué categoría de riesgo clasifica el regulador a un sistema como ChatGPT, Claude o Gemini? ¿Riesgo mínimo, porque a veces sugiere música? ¿Alto riesgo, porque puede ser utilizado para elaborar informes periciales o apoyar decisiones administrativas? ¿Riesgo inaceptable, porque también podría ser instrumentalizado para influir en conductas o hábitos de personas? La respuesta es que el reglamento no lo sabe, porque fue diseñado para clasificar herramientas de función fija, no tecnologías de función abierta.

1.3.2. Estados Unidos

En Estados Unidos, la regulación de la IA se caracteriza por la ausencia de un marco legislativo federal integral que regule de manera uniforme el desarrollo, implementación y uso de estos sistemas. En su lugar, el sistema estadounidense ha evolucionado a través de un modelo regulatorio fragmentado, compuesto por la aplicación de normas preexistentes, iniciativas administrativas del poder ejecutivo y una creciente producción normativa a nivel estatal, que refleja una tradición jurídica que privilegia la innovación tecnológica y la competitividad económica, procurando evitar regulaciones prematuras que puedan obstaculizar el desarrollo de nuevas tecnologías.

Recién en la presidencia de Biden se realizó el primer intento regulatorio, promulgando el 2022 el *Al Bill of Rights*. Es un documento flexible, no vinculante, que establece la política de Estados Unidos sobre el desarrollo de la IA. Se centra en aspectos que giran en torno a la existencia de los ciudadanos, pero resulta insuficiente. Sin embargo, el año 2023 esta norma se complementa de forma omnicomprensiva, mediante la Orden Ejecutiva 14110 , sobre el Desarrollo y Uso Seguro, Confiable y Protegido de la IA. Es una política pública que se sostiene en principios tales como la seguridad y protección; privacidad y libertades civiles; equidad y no discriminación; innovación y competencia; protección de la fuerza laboral; seguridad nacional y liderazgo global; y el gobierno como usuario modelo.

En particular, me parece relevante destacar que la equidad y no discriminación implicaba que los modelos debían estar dispuestos a pasar por un proceso de auditorías en materia de sesgo algorítmico, estableciendo también, los mecanismos para combatirlos con especial protección en las comunidades marginadas. Además, el gobierno debe implementar la IA en la administración pública conforme a las mejores prácticas, de manera segura y transparente.

A pesar de que esta normativa se encontraba bien encaminada, era un proyecto en constante construcción. Pero fue derogada el 20 de enero de 2025 por el gobierno del Presidente Trump, que luego, seis meses después, presentaría un nuevo documento denominado *Winning the Race. America`s AI Action Plan*, fundado en el aceleramiento de la innovación, la construcción de infraestructura estadounidense de IA, y, en definitiva, ser pioneros en diplomacia y seguridad internacional en la materia (*The White House*, 2025, p. 9).

Al ser uno de los principales desarrolladores de estos sistemas, se esperaría una legislación robusta. Sin embargo, el documento se limita a establecer los principios fundamentales, eliminando trabas burocráticas, promoviendo la inversión de datos, infraestructura energética, investigación en IA e integración de esta en el entorno administrativo del gobierno. Es una política que pareciera priorizar la velocidad para acelerar el ritmo en el desarrollo de estas tecnologías, pero que no se preocupa de dilemas regulatorios importantes como la responsabilidad civil, implicando incluso, un retroceso en materia de sesgo algorítmico.

En relación con lo anterior, ante la Corte del Distrito Medio de Florida se ventila actualmente un litigio de particular relevancia, dirigido en contra de las empresas *Character Technologies Inc., Google LLC y Alphabet Inc.*, cuyo objeto es la presunta responsabilidad civil derivada del diseño defectuoso de un sistema de IA generativa que, según se alega, habría inducido al suicidio a un joven de catorce años. El caso no es solo una tragedia humana, sino también un *test* judicial sin precedentes respecto de los límites de la responsabilidad civil. Según reportó Natallie Roche y Kashmir Hill en *The New York Times* (29 de octubre de 2025), la demanda se estructura sobre la base de la responsabilidad estricta por productos defectuosos o peligrosos, imputando a las demandadas haber dotado deliberadamente al *chatbot* de cualidades antropomórficas destinadas a difuminar la frontera entre la ficción y la realidad.

La demandante solicita indemnización de perjuicios, sosteniendo que este diseño no fue accidental ni neutral, sino que buscaba captar de forma adictiva la voluntad de usuarios vulnerables, sin que se implementaran salvaguardas de seguridad suficientes ni advertencias sobre los riesgos que su uso prolongado podría representar para la salud mental, fundándose en la omisión de estándares mínimos de debida diligencia en la comercialización del *software*. Sobre este punto, las palabras de los padres del joven fallecido adquieren una dimensión que ningún texto académico puede ignorar; señalaron que, sin el *chatbot*, su hijo quizás seguiría sufriendo y necesitando ayuda, pero al menos estaría vivo. Pocas afirmaciones sintetizan con tanta crudeza la magnitud del bien jurídico en juego. En conjunto, los cargos configuran un cuadro de imputación que desborda la lógica tradicional de la responsabilidad del fabricante y obliga a preguntarse si las categorías existentes son suficientes para capturar la complejidad del daño tecnológico.

1.3.3. Chile

La experiencia comparada permite identificar una tendencia entre los distintos modelos regulatorios que los Estados han desarrollado frente a la IA. La búsqueda de un equilibrio entre el fomento de la innovación tecnológica y la protección efectiva de los derechos fundamentales. No se trata de una tensión nueva para el derecho; es, más bien, la versión contemporánea de una pregunta que el ordenamiento jurídico ha debido enfrentar cada vez que una tecnología disruptiva ha reconfigurado las condiciones bajo las cuales se producen los daños y se distribuyen los riesgos sociales. Lo que sí resulta preocupante es la velocidad con que esa pregunta se vuelve urgente y la profundidad con que interpela categorías jurídicas que parecían consolidadas.

Los desafíos regulatorios no se agotan en la definición conceptual de la IA, sino que se concentran en cuestiones bastante más complejas, como la atribución de responsabilidades en cadenas de valor tecnológicas fragmentadas, la exigencia de transparencia sobre procesos algorítmicos que por diseño resultan opacos, y la capacidad institucional efectiva para supervisar sistemas cuyo funcionamiento interno escapa a los métodos de fiscalización tradicionales. Estas señales del derecho comparado constituyen un primer antecedente para el caso chileno, en que actualmente se desarrolla un proceso de consolidación normativa que intenta incorporar tales preocupaciones dentro de un marco jurídico propio.

La regulación de la IA en Chile se encuentra, hoy por hoy, en una etapa de transición. El país ha comenzado a evolucionar desde un sistema articulado fundamentalmente sobre políticas públicas de carácter orientador hacia la construcción de una normativa legal vinculante, con todo lo que aquello implica en términos de certeza jurídica, exigibilidad y consecuencias ante el incumplimiento.

Hasta la fecha, Chile no cuenta con una ley específica vigente sobre IA, de modo que el marco regulatorio ha descansado principalmente en la Política Nacional de IA, publicada originalmente en 2021 y actualizada en 2024 por el Ministerio de Ciencia, Tecnología, Conocimiento e Innovación. Este instrumento establece lineamientos estratégicos y éticos para el desarrollo de la IA en el país, articulados en torno a principios de inclusión, responsabilidad,

transparencia y desarrollo sostenible (Ministerio de Ciencia, Tecnología, Conocimiento e Innovación, 2024).

Es importante detenerse en lo que significa construir la gobernanza de una tecnología sobre la base de políticas públicas no vinculantes que orientan, recomiendan y expresan las prioridades del Estado, pero no es obligatorio. En un entorno tecnológico en que las decisiones de diseño de un sistema algorítmico pueden afectar simultáneamente la salud, la libertad, el empleo y la dignidad de miles de personas, la diferencia entre una recomendación ética y una norma jurídica exigible es la diferencia entre un sistema de gobernanza con consecuencias reales y uno que depende exclusivamente de la buena voluntad de sus destinatarios. Chile ha reconocido esta limitación, y el proceso legislativo en curso es la respuesta institucional a esa constatación.

En particular, el eje central del proceso regulatorio chileno se encuentra en el Proyecto de Ley que Regula los Usos de la IA (Boletín N.º 16821-19), ingresado al Congreso Nacional en mayo de 2024 y que, a la fecha, se encuentra en su segundo trámite constitucional ante el Senado de la República (Congreso Nacional de Chile, 2024). La sola trayectoria legislativa del proyecto es ya ilustrativa de los tiempos que demanda traducir una política tecnológica en derecho positivo, en un contexto en que la tecnología que se pretende regular no espera.

Siguiendo las tendencias regulatorias observadas en otras jurisdicciones, en especial la Unión Europea, el proyecto adopta un modelo de regulación basado en la gestión del riesgo, clasificando los sistemas de IA en cuatro categorías que van de menor a mayor intensidad regulatoria. En primer lugar, los sistemas de riesgo inaceptable, cuyo uso queda directamente prohibido por resultar incompatible con los derechos fundamentales. En segundo lugar, los sistemas de alto riesgo, que incluyen aplicaciones capaces de afectar la salud, la seguridad o los derechos de las personas, y que quedan sujetos a obligaciones más exigentes. En tercer lugar, los sistemas de riesgo limitado, sometidos principalmente a deberes de transparencia orientados a garantizar que los usuarios sepan que están interactuando con un sistema automatizado. Finalmente, los sistemas de riesgo mínimo o inexistente, que no quedan sujetos a obligaciones regulatorias especiales (Araya Paz, 2020, pp. 257–290). Esta arquitectura permite vincular la intensidad de las cargas jurídicas al nivel de impacto potencial que cada sistema puede generar sobre las personas, lo que desde el punto de vista de la proporcionalidad regulatoria es un diseño razonable, aunque no exento de problemas.

Según señalan Amunátegui, Aránguiz y Jara (2025, pp. 15-21), si bien el Reglamento europeo ha operado como fuente principal de inspiración de la propuesta chilena, esa recepción no puede efectuarse sin una debida contextualización a la realidad local, considerando que Chile representa una industria de menor escala, con recursos considerablemente inferiores a los de los países líderes en IA, y que carece de la base regulatoria previa en protección de datos y ciberseguridad de la que sí dispone la Unión Europea. Por esta razón, Chile requiere una regulación más simple, que concentre sus esfuerzos en los ámbitos de mayor relevancia y en la viabilidad efectiva de su fiscalización, antes que en la réplica exhaustiva del modelo comunitario.

Con todo, existen aspectos positivos en la iniciativa legislativa, entre ellos su potencial para atraer inversión tecnológica, su alineación con los estándares internacionales y la incorporación de espacios controlados de prueba o *sandboxes* regulatorios (Amunátegui et al., 2025, pp. 15-21). Sin embargo, estos méritos conviven con falencias significativas, identificadas a partir de la revisión de más de treinta y cinco intervenciones recibidas durante la discusión parlamentaria del Boletín 16.821-19.

Entre dichas falencias destacan la amplitud excesiva de la definición de sistema de IA, que termina sometiendo a un mismo estatuto normativo a tecnologías de naturaleza muy distinta, generando simultáneamente sobrerregulación y subregulación; la vaguedad de conceptos como riesgo inaceptable y alto riesgo; y la existencia de principios, como la transparencia, la explicabilidad y la rendición de cuentas, que carecen de contenido material suficiente para ser aplicados en la práctica (Amunátegui et al., 2025, pp. 15-21).

La calidad del corpus utilizado en la construcción de los modelos no debiera abordarse desde una perspectiva ética o principialista, sino desde la estricta licitud jurídica del dato, cuestión que ya encuentra cauce regulatorio en la Ley de Protección de Datos Personales y en la Ley Antidiscriminación, de manera que exigir un adicional alineamiento ético en la nueva normativa generaría una duplicidad innecesaria, pues permite reconducir un problema que suele tratarse en términos excesivamente abstractos hacia estándares de cumplimiento normativo ya existentes y verificables. (Amunátegui et al., 2025, pp. 15-21).

Junto con la clasificación de riesgos, el proyecto de ley de IA introduce la noción de operadores como categoría comprensiva de los proveedores, implementadores, importadores y

distribuidores de sistemas de IA. Sobre estos actores recaen distintas obligaciones jurídicas dependiendo del nivel de riesgo del sistema que desarrollen o utilicen, configurando así un modelo de responsabilidad distribuida a lo largo de la cadena de valor tecnológica. La lógica es atendible, si la producción, comercialización y despliegue de un sistema de IA involucra a múltiples actores con distintos grados de control sobre su funcionamiento, tiene sentido que la regulación reconozca esa complejidad en lugar de concentrar toda la imputación en un único sujeto. El desafío, como se verá, está en que esa distribución de responsabilidades no se convierta en una disipación de las mismas.

A pesar de estas tensiones, el proceso regulatorio chileno evidencia un esfuerzo deliberado por alinear su gobernanza tecnológica con estándares internacionales, particularmente aquellos promovidos por la Unión Europea y la OCDE. Las modificaciones introducidas en la actualización de la política pública de IA reflejan una convergencia con estos marcos regulatorios globales, especialmente en materia de ética algorítmica, transparencia y evaluación de riesgos (Tellez Tejada, 2025). Esta orientación es comprensible desde el punto de vista de la inserción de Chile en el ecosistema tecnológico global. En efecto, operar bajo estándares reconocidos internacionalmente facilita la interoperabilidad regulatoria, reduce los costos de cumplimiento para empresas que operan en múltiples jurisdicciones y genera señales de confianza para los inversionistas tecnológicos.

Sin embargo, como sugieren Torres Mac-Pherson et al. (2025, pp. 26-37), el verdadero desafío institucional no radica únicamente en adoptar estándares internacionales, sino en adaptar las categorías tradicionales del derecho chileno a la naturaleza dinámica, interconectada y en ocasiones imprevisible de los sistemas de IA. Importar una categoría regulatoria sin atender a cómo se articula con el resto del ordenamiento jurídico nacional puede generar inconsistencias que terminen por debilitar la eficacia del marco normativo en su conjunto. El derecho comparado es una fuente de inspiración valiosa, pero no es un catálogo de soluciones trasplantables sin adaptación.

El proceso de regulación de la IA en Chile se encuentra, en definitiva, ante una disyuntiva estructural que no admite soluciones simples. Por una parte, debe promover un entorno normativo que facilite el desarrollo tecnológico y la innovación, reconociendo que la IA representa una oportunidad de transformación económica y social que el país no puede ignorar. Por otra, debe

garantizar que la expansión de estas tecnologías no genere espacios de desprotección jurídica para las personas afectadas por fallos algorítmicos o por decisiones automatizadas que incidan en sus derechos fundamentales. Ambos objetivos son legítimos, y la tensión entre ellos no se resuelve eligiendo uno sobre el otro, sino construyendo un marco regulatorio suficientemente sofisticado para sostenerlos simultáneamente.

Al examinar el marco regulatorio europeo y contrastarlo con la realidad chilena, me parece indispensable advertir que ambos ordenamientos parten de posiciones estructuralmente distintas frente a la IA, lo que impide trasladar sin más el modelo de la Unión Europea al proyecto de ley que actualmente se tramita en Chile. La Unión Europea ha construido su Reglamento desde la posición de un actor que, en gran medida, produce, desarrolla e investiga estas tecnologías dentro de su propio territorio. De ahí que su regulación se concentre, principalmente, en imponer obligaciones previas y estándares de diligencia a quienes diseñan, entrenan y ponen en circulación los sistemas de IA, exigiéndoles una gestión de riesgos desde la etapa de concepción del producto.

Chile, en cambio, no ocupa ese lugar en la cadena tecnológica. Nuestro país no es, salvo excepciones puntuales, un polo de desarrollo o innovación en esta materia, sino fundamentalmente un receptor de sistemas diseñados, entrenados y comercializados en otras jurisdicciones. Esta diferencia determina consecuencias jurídicas de fondo.

Si se aprueba el proyecto de ley chileno, que replica el Reglamento UE, corre el riesgo de regular una realidad que, en los hechos, escapa en gran medida a su jurisdicción. Buena parte de los sistemas que se utilizarán en Chile habrán sido diseñados y entrenados en el extranjero, conforme a estándares, datos y decisiones de diseño adoptadas por empresas que no necesariamente tienen presencia física ni patrimonio perseguible dentro del territorio nacional.

La eficacia de la futura legislación chilena dependerá, en última instancia, de su capacidad para adaptarse a una tecnología que desafía los conceptos tradicionales de control, previsibilidad y responsabilidad en el derecho contemporáneo. Una ley que regule bien la IA de 2024 pero que no prevea mecanismos de actualización frente a la IA de 2030 habrá repetido, en escala nacional, el mismo error que ya cometió la Unión Europea a escala continental. Y esa es, quizás, la lección más importante que el derecho comparado tiene para ofrecerle al proceso regulatorio chileno, que la

mayor amenaza para una buena regulación tecnológica no es la imperfección inicial, sino la rigidez frente al cambio.

1.4. Conclusiones preliminares

El análisis del derecho comparado no debería leerse como un inventario de modelos a imitar, sino como un registro de errores que todavía es posible no cometer. Esa es, quizás, la mayor utilidad que el estudio del derecho extranjero puede ofrecer a un proceso legislativo en curso. A partir de lo anterior, extraigo las siguientes reflexiones:

1. El recorrido histórico y conceptual me permite afirmar que la IA no constituye un fenómeno jurídicamente nuevo en su lógica, sino en su escala y en su opacidad. Los problemas que hoy enfrenta el derecho de daños, la dificultad para identificar una conducta negligente, para reconstruir un nexo causal o para individualizar al sujeto responsable, no son inéditos; lo inédito es que estos problemas se presentan ahora multiplicados por sistemas que aprenden, que se adaptan y que, en ciertos casos, ni sus propios creadores logran explicar del todo. El problema de la caja negra, más que una curiosidad técnica, es el síntoma jurídico central de esta transformación; un daño cuyo origen resulta difícil de trazar desafía, en su raíz, la estructura probatoria sobre la que se estructura la responsabilidad civil clásica.
2. El aporte de Rumelhart, más allá de su valor técnico, me parece revelador para el problema de imputación que ocupa esta tesis. Al estructurar el algoritmo en capas intermedias, entre la entrada y la salida, sentó las bases de una arquitectura que, en principio, permite descomponer la decisión final en las distintas etapas de análisis que la fueron construyendo. Esto haría posible el desarrollo de técnicas de IA explicable, orientadas a abrir la caja negra y reconstruir el itinerario que condujo a un resultado determinado. En consecuencia, si la propia arquitectura del sistema permite, aunque sea parcialmente, trazar el camino recorrido entre el dato de entrada y la decisión final, entonces la opacidad algorítmica deja de ser una barrera insalvable para la causalidad y se convierte en un problema de diseño técnico susceptible de exigencia jurídica. Me parece, en consecuencia, que el principio de

explicabilidad no es solo una aspiración ética, sino una condición de posibilidad para la propia responsabilidad civil, en la medida en que permite reconstruir, con algún grado de razonabilidad, aquello que la estructura en capas mostró que sí puede ser una alternativa.

3. El examen del derecho comparado me lleva a sostener que existe una tendencia regulatoria común, orientada hacia modelos basados en la gestión de riesgo, pero que dicha convergencia formal esconde diferencias sustanciales en el modo de enfrentar la responsabilidad civil. La Unión Europea articula su Reglamento desde una posición de productor tecnológico, lo que explica que su arquitectura normativa se concentre en imponer obligaciones previas y estándares de diligencia a lo largo de todo el ciclo de vida del sistema, desde su experimentación hasta su despliegue efectivo. Estados Unidos, en cambio, revela las tensiones de un modelo fragmentado y oscilante, en que la ausencia de una legislación federal estable traslada buena parte de la definición de los estándares de responsabilidad a la vía judicial. Por su parte, Chile se encuentra en un proceso de transición en que la política pública antecede a la norma vinculante, evidenciando que la construcción de un marco regulatorio es un proceso de consolidación institucional.

4. El error más significativo que Chile podría cometer no es tanto legislar tarde, sino legislar copiando un modelo pensado para una realidad productiva distinta a la propia. Chile interviene en esta cadena fundamentalmente como consumidor. Trasladar sin matices un régimen centrado en las obligaciones del proveedor a un país en que la mayoría de los sistemas de IA son diseñados, entrenados y comercializados en el extranjero corre el riesgo de producir una tutela meramente declarativa, en la que la víctima carezca de un sujeto identificable y patrimonialmente exigible dentro del territorio nacional. De ahí que, en mi opinión, el centro de gravedad regulatorio deba desplazarse hacia quienes efectivamente despliegan, distribuyen e importan estos sistemas en Chile, pues son ellos quienes se encuentran al alcance real de nuestros tribunales y órganos de fiscalización.

5. El análisis del proceso regulatorio europeo, marcado por el desfase entre una arquitectura normativa concebida en 2015 y una tecnología que para 2024 ya la había superado, me deja una lección que trasciende el caso comunitario y que Chile debiera internalizar desde ya: la principal amenaza para una buena regulación de la IA no es la imperfección con que nace, sino la rigidez con que resiste el cambio. Un marco jurídico que fije categorías cerradas para una tecnología en constante transformación corre el riesgo de envejecer antes de entrar

plenamente en vigencia. Por ello, considero que el desafío que enfrenta Chile no se agota en aprobar una ley sobre IA, sino en dotarla de mecanismos de actualización y de una comprensión clara de su posición como país consumidor, condiciones ambas indispensables para que la responsabilidad civil por daños derivados de estos sistemas sea, en la práctica, algo más que una promesa normativa.

CAPÍTULO 2: EL DERECHO DE DAÑOS ANTE LA INTELIGENCIA ARTIFICIAL

2.1. La función preventiva de la responsabilidad civil en el contexto tecnológico

La expansión de la IA en múltiples ámbitos de la vida social, y particularmente en el ejercicio de profesiones jurídicas y funciones públicas, obliga a reconsiderar que la responsabilidad civil es, ante todo, un mecanismo resarcitorio. Esa lectura es correcta, pero incompleta. En contextos caracterizados por la creciente automatización de procesos decisionales, la responsabilidad civil actúa como instrumento de regulación preventiva del riesgo tecnológico. Su función no es solo reparar el daño una vez producido, sino establecer estándares de conducta *ex ante* para quienes diseñan, implementan o utilizan sistemas de IA, incentivando la adopción de medidas de diligencia razonables destinadas a impedir esos daños.

La responsabilidad civil puede ser entendida como una técnica de prevención general, en la medida en que sus reglas emiten señales de cuidado que operan como un mercado autorregulado de incentivos. Bajo este modelo, quien desarrolla o despliega un sistema de IA conoce, de antemano, las condiciones que determinan su deber de indemnizar a las víctimas de los daños que dicha actividad pueda ocasionar. Se establece así, implícitamente, un desincentivo al descuido, puesto que el operador deberá asumir el costo de los daños subsecuentes, y un incentivo correlativo a incurrir en gastos de prevención (Barros, 2020, pp. 47-48).

Este razonamiento permite sostener que un desarrollador o proveedor que internaliza el costo esperado de una eventual indemnización tiene razones económicas suficientes para invertir en auditorías de sesgo algorítmico, en protocolos de seguridad y en mecanismos efectivos de supervisión humana, con el objeto de reducir la probabilidad de que el sistema genere un daño. La responsabilidad civil no prescribe, en este esquema, una conducta técnica determinada, sino que fija el costo asociado a no adoptarla, dejando a criterio del operador la forma en que decide prevenir el riesgo (Barros, 2020, pp. 47-48).

En consecuencia, las reglas de responsabilidad civil operan como directivas de prevención general dirigidas a todo creador de riesgo, estableciendo una regla de atribución de los costos derivados de los daños asociados a su actividad. Aplicado a la cadena de valor de la IA, ello significa que desarrolladores, proveedores y responsables del despliegue quedan sujetos a un mismo trasfondo normativo, que orienta sus decisiones sin que el Estado deba intervenir de manera casuística en el diseño técnico de cada sistema.

Cabe agregar que el ordenamiento civil actúa como un mecanismo descentralizado de control, toda vez que es la propia víctima del daño algorítmico quien debe ejercer las acciones pertinentes para hacer valer su derecho a ser indemnizada. La función del Estado se circunscribe, entonces, a establecer las reglas de responsabilidad y a proveer los medios procesales necesarios para su ejercicio, configurándose de este modo un régimen descentralizado de incentivos que propende a un óptimo de bienestar general (Barros, 2020, pp. 47-48).

En definitiva, esta función no busca suprimir por completo la posibilidad de que ocurran daños, sino alcanzar un grado razonable de prevención, que pondere tanto los costos que una determinada actividad genera como los beneficios sociales que de ella se derivan. Las reglas de responsabilidad civil, en tanto establecen incentivos de mercado asociados a la imposición de sanciones indemnizatorias, permiten mantener esos costos dentro de niveles socialmente tolerables. Desde esta perspectiva económica, las normas de responsabilidad no se valoran, entonces, por su capacidad de erradicar el riesgo, sino por su eficiencia dentro de un balance general de bienestar (Barros, 2020, p. 46).

Con todo, cuando los costos de acceder a la información, ejercer supervisión o llevar adelante una negociación resultan excesivamente altos, el ordenamiento puede recurrir a instrumentos de intervención más intensos, como las llamadas reglas de inalienabilidad, que según Calabresi y Melamed (1972, pp. 1089-1128) habilitarían a restringir o incluso prohibir determinadas aplicaciones tecnológicas. Esta lógica se refleja en distintos marcos regulatorios comparados, respecto al uso de sistemas de IA calificados como de riesgo prohibido, precisamente por su potencial incidencia en ámbitos sensibles. En esos contextos, la intervención estatal se justifica en la necesidad de preservar bienes jurídicos valiosos y evitar que una confianza excesiva en la autonomía tecnológica comprometa la legitimidad de los procesos institucionales.

Finalmente, reconocer la función preventiva no significa desplazar ni relativizar su función primaria y, en definitiva, más antigua de la reparación del daño ya causado. Por relevantes que sean los incentivos que el sistema genera hacia el futuro, si existe una víctima concreta cuyo patrimonio, salud o derechos han sido efectivamente afectados por el funcionamiento de un sistema de IA, el ordenamiento debe dar una respuesta reparatoria adecuada. En este punto adquiere sentido el principio de la reparación integral del daño, conforme al cual la víctima tiene derecho a ser restituida a la situación patrimonial en que se encontraría de no haber sufrido el perjuicio. A continuación, me propongo examinar de qué manera este principio, de larga tradición en el derecho de daños, se proyecta sobre los perjuicios derivados del uso de sistemas de IA, y qué particularidades presenta su aplicación frente a daños que, con frecuencia, resultan de procesos algorítmicos complejos, opacos y de difícil cuantificación.

2.2. El principio de la reparación integral de los daños derivados en uso de IA

El artículo 2329 del Código Civil dispone que "todo daño atribuible a la malicia o negligencia de otra persona debe ser reparado por esta". Si bien esta norma tiene por finalidad principal establecer una presunción de responsabilidad por el hecho propio, la jurisprudencia también ha reconocido que se desprende el principio de la reparación integral del daño, conforme al cual todo perjuicio debe ser indemnizado en su totalidad (Barros, 2020, pp. 255 y 256). Este principio constituye, a mi juicio, el criterio más general para determinar el alcance de la indemnización, y me parece que resulta especialmente relevante cuando se trata de daños causados por sistemas de IA, cuya magnitud puede ser considerable y de difícil previsión inicial.

De acuerdo con este principio, la reparación busca colocar a la víctima en la misma situación patrimonial en que se encontraría de no haber sufrido el daño ocasionado por el hecho del demandado (Barros, 2020, pp. 255 y 256). Aplicado a los perjuicios derivados del funcionamiento de un sistema de IA, esto significa que la indemnización no puede limitarse arbitrariamente en atención a la complejidad técnica del sistema que produjo el daño, sino que debe cubrir la totalidad del perjuicio efectivamente sufrido por la víctima.

De lo anterior se siguen dos consecuencias que considero relevantes. La primera es el deber de reparar el total del daño causado. La segunda, derivada de la anterior, es que la extensión de la reparación no depende del grado de culpa del demandado. Esto implica que, si se incurre en una negligencia al utilizar o supervisar una herramienta de IA, pero de dicha negligencia se sigue un daño considerable, igualmente deberá responder por la totalidad del perjuicio, sin que pueda invocar la levedad de su culpa para reducir el monto indemnizatorio.

Esta consecuencia ha sido objeto de crítica, incluso en el ámbito general de la responsabilidad por culpa, pues puede generar una carga indemnizatoria desproporcionada respecto de la entidad de la negligencia cometida. Sin embargo, el principio de reparación integral se vincula estrechamente con la idea de justicia correctiva, según la cual quien es responsable de un daño debe reparar el entuerto provocado, con independencia del reproche que merezca su conducta (Barros, 2020, pp. 255 y 256). Apartarse de este principio en el ámbito de la IA podría, a mi juicio, debilitar la función reparatoria que el sistema de responsabilidad civil está llamado a cumplir.

Con todo, reconozco que el propio ordenamiento contempla mecanismos correctivos frente a resultados inequitativos, del mismo modo que el principio de reparación integral admite límites en ciertos estatutos especiales de responsabilidad, particularmente en regímenes de responsabilidad estricta. Estos correctivos resultan, a mi parecer, especialmente pertinentes a la hora de diseñar un régimen de responsabilidad civil aplicable a los daños derivados del uso de la IA, en la medida en que permiten conciliar la exigencia de reparación integral con la necesidad de no imponer cargas indemnizatorias desproporcionadas a quienes participan en el desarrollo o utilización de estas tecnologías.

Ahora bien, la vigencia de este principio no puede examinarse en abstracto, sino que debe confrontarse con las particularidades técnicas que presentan los sistemas cuyo daño se pretende reparar. Me parece pertinente detenerme, entonces, en dos características de los sistemas de IA que la doctrina especializada identifica como los principales factores de tensión frente a las reglas tradicionales de la responsabilidad civil: la autonomía y la opacidad. Se trata de elementos que inciden directamente en la posibilidad de hacer operativo, en la práctica, el derecho de la víctima a una reparación integral.

La autonomía se traduce en la capacidad de estos sistemas para adoptar decisiones que no fueron predeterminadas ni, en muchos casos, previsibles para quienes intervinieron en su desarrollo u operación (Ayo Ferrándiz, et al., 2025, pp. 29-56). Esta característica complica, a mi juicio, la atribución de responsabilidad conforme a las reglas clásicas de imputación subjetiva, pues dichas reglas presuponen que el sujeto responsable pudo prever, y por tanto evitar, la conducta generadora del daño. Asimismo, no puedo dejar de considerar que la singularidad que distingue a la IA de toda tecnología precedente radica precisamente en que es la primera creación humana capaz de evolucionar por sí misma, sin intervención directa de quien la diseñó.

La opacidad, por su parte, alude a la dificultad, cuando no imposibilidad, de comprender el funcionamiento interno de un sistema de IA y, en particular, de explicar las razones que llevaron a la adopción de una decisión determinada. Esta característica, conocida como el efecto de la caja negra, incide de manera directa en el juicio de causalidad, tanto fáctica como jurídica, que resulta indispensable para configurar la responsabilidad extracontractual (Ayo Ferrándiz, et al., 2025, pp. 29-56). Si no es posible reconstruir el proceso que condujo al resultado dañoso, tampoco resulta sencillo establecer el vínculo causal que el sistema de responsabilidad civil exige entre la conducta de un sujeto determinado y el perjuicio finalmente producido.

Cabe recordar, en este punto, lo señalado en el primer capítulo respecto de la contribución de Rumelhart al desarrollo de las redes neuronales, particularmente en lo que atañe a la incorporación de capas intermedias entre los mecanismos de entrada y salida del algoritmo. Esta posibilidad, que hoy se desarrolla bajo la denominación de IA explicable, no resuelve por sí sola las dificultades probatorias que la opacidad algorítmica plantea al derecho de daños, pero sí permite sostener que dicha opacidad no es una condición insuperable de estos sistemas, sino, en buena medida, una decisión de diseño técnico. De ahí que considere razonable exigir, como estándar de diligencia aplicable a quienes desarrollan y despliegan sistemas de IA, la adopción de arquitecturas y mecanismos que favorezcan la trazabilidad de sus decisiones, pues solo así el principio de explicabilidad puede operar como un verdadero instrumento al servicio de la reconstrucción del nexo causal, que permita, en definitiva, facilitar la dificultad probatoria de la víctima.

Delineadas estas consideraciones sobre las funciones de la responsabilidad civil frente a los riesgos de la IA, me parece necesario avanzar hacia un examen de las posibles formas de imputación jurídica de los daños que estas tecnologías pueden generar. En particular, corresponde

analizar cómo los distintos regímenes tradicionales de responsabilidad civil podrían proyectarse sobre estos perjuicios, atendiendo a las tensiones que emergen entre las categorías clásicas del derecho de daños y las características propias de estos sistemas. Ese análisis permitirá identificar las principales dificultades que plantea la atribución de responsabilidad en este ámbito y, con ello, los desafíos que enfrenta el derecho para adaptar sus herramientas al entorno tecnológico.

2.3. Aplicación de los sistemas de atribución de responsabilidad civil

Las personas estamos normalmente expuestas a una pluralidad de riesgos por el solo hecho de existir en sociedad, exposición que se ha intensificado con la introducción de la IA. Basta pensar en el postulante a un empleo que es descartado por un sistema de selección automatizado que reproduce sesgos discriminatorios, en el paciente cuyo diagnóstico se ve alterado por un algoritmo de apoyo clínico que procesó incorrectamente sus antecedentes, o en el litigante cuyo caso se ve perjudicado por un informe jurídico elaborado con una herramienta de IA que incurrió en un error.

Frente a esta pluralidad de riesgos, quien sufre un daño debe soportarlo, salvo que exista un fundamento jurídico que permita imputarlo a otro sujeto. El caso del usuario de un *chatbot* de uso general resulta ilustrativo de esta regla; quien decide interactuar con estos sistemas, conociendo su carácter experimental y las advertencias sobre sus limitaciones, asume como propio el riesgo de recibir una respuesta imprecisa o inexacta, de manera que el daño derivado de ese uso no puede imputarse al desarrollador, salvo que medie una infracción específica.

Distinta es, sin embargo, la situación cuando el daño no encuentra su origen en una decisión libremente asumida por la víctima, sino en la conducta de un tercero que introduce o agrava el riesgo del cual ese perjuicio deriva. En un caso así, el ordenamiento jurídico permite trasladar el peso económico del daño desde la víctima hacia el tercero cuya conducta lo provocó. La responsabilidad civil es, precisamente, el mecanismo que establece los criterios normativos para efectuar esa imputación. El sistema articula un conjunto de requisitos que habilitan la atribución jurídica del daño a quien lo causó (Barros Bourie, 2020, pp. 24-31).

Aceptado entonces que un tercero ha intervenido causalmente en la producción de un daño, la pregunta es si corresponde que la víctima sea indemnizada por él. Como advierte Barros (2020, pp. 24-31), la respuesta depende de las condiciones que el derecho establezca para que esa responsabilidad proceda. De ahí surge distinguir si basta la mera intervención causal del tercero, o si es necesario además que su conducta resulte jurídicamente reprochable. De esta distinción emergen los dos grandes sistemas de atribución: el régimen de responsabilidad estricta u objetiva, que exige acreditar solo un vínculo causal entre la actividad y el daño, trasladando al agente los riesgos inherentes a esa actividad; y el régimen de responsabilidad por culpa, que exige demostrar que el autor actuó con negligencia, imprudencia o mala fe.

Corresponde, entonces, detenerse en el primero de estos regímenes. La responsabilidad objetiva constituye un modelo de imputación que prescinde de la exigencia de probar culpa, fundamentándose en el riesgo creado por la actividad o el producto (Barros Bourie, 2020). Su atractivo en el contexto de la IA radica en que facilita la posición de la víctima, eximiéndola de la prueba del factor de atribución o imputabilidad. Según Atienza (2022, pp. 67-68), quienes reconocen la existencia de una IA fuerte, con capacidad de aprendizaje genuino y autonomía decisional, tienden a abogar por este sistema, argumentando que las características de opacidad, complejidad y autonomía hacen imposible imputar daños mediante el filtro de la culpa.

Para comprender por qué la culpa resulta insuficiente en estos casos, es necesario recordar su lógica de funcionamiento. La culpa constituye, tradicionalmente, una herramienta de distribución de riesgos que opera mediante un estándar de conducta socialmente aceptado, a partir del cual se determina si el agente incurrió o no en negligencia. Presupone que el riesgo generado por la actividad corresponda, precisamente, a aquel que el estándar de diligencia exigible está en condiciones de prevenir. Cuando el riesgo excede ese umbral, la responsabilidad objetiva se presenta como el fundamento más adecuado, en la medida en que este régimen se justifica frente a la gestión de riesgos extraordinarios o anormales, categoría que comprende tanto aquellos supuestos en que existe una probabilidad especialmente alta de que el daño se produzca, como aquellos en que, de materializarse, sus consecuencias resultarían catastróficas Miranda Valdebenito, 2025, pp. 103-129).

Con el objeto de precisar el contenido del concepto de peligrosidad anormal, resulta pertinente el criterio propuesto por Canaris y Larenz (1994, pp. 605 y ss.), quienes sugieren

examinar la relación existente entre el agente y el riesgo, atendiendo especialmente a si aquel se encuentra en posición de dominarlo, en el sentido de que, no obstante desplegar las medidas destinadas a evitar el daño, este pueda igualmente producirse. Esta hipótesis da lugar a los denominados daños residuales, categoría que permite ilustrar con particular claridad la relación entre el agente artificial y el ser humano en el momento inmediatamente posterior a su actuación.

En estos supuestos, la decisión adoptada en el instante previo a la producción del daño escapa al control humano, de manera que las medidas de prevención adoptadas con anterioridad resultan insuficientes. Así ocurre, por ejemplo, con un vehículo de conducción autónoma programado para evitar colisiones; si un peatón aparece intempestivamente en la calzada, el sistema deberá optar entre atropellarlo o colisionar contra otro vehículo. Este ejemplo permite constatar que, aun cuando se hayan adoptado las medidas de prevención pertinentes, el riesgo puede materializarse de todas formas (Miranda Valdebenito, 2025, p. 105).

A partir de lo anterior, concluyo que la utilización de sistemas de IA fuerte introduce, en sí misma, un riesgo anormal respecto del tráfico social habitual. Así se aprecia, por ejemplo, en el ámbito de la conducción autónoma, en que un conductor humano puede no estar en condiciones de advertir que el vehículo que circula delante suyo es operado por un sistema automatizado. La sola presencia de un vehículo de estas características implica, entonces, un riesgo anormal, que puede derivar tanto de una toma de decisión inesperada por parte del sistema como de una falla en su proceso decisional, en los términos ya ejemplificados. Este fundamento permite abordar adecuadamente el problema de la distribución de riesgos que la IA genera, habilitando la atribución objetiva de responsabilidad en aquellos supuestos que exceden las dinámicas habituales y, por tanto, previsibles para los miembros de la comunidad (Miranda Valdebenito, 2025, pp. 103-129).

Para Brüggemeier (2006, p. 114), el concepto central de la responsabilidad objetiva radica en el denominado ámbito de riesgo cubierto, en la medida en que este régimen busca determinar si resulta justo y razonable que un riesgo específico quede amparado por la protección que la responsabilidad objetiva otorga. En un sentido análogo, Prosser y Keeton (1984, p. 562) sostienen que la responsabilidad objetiva se circunscribe, por regla general, a aquellas consecuencias que caen dentro del riesgo extraordinario cuya existencia justifica precisamente ese régimen especial de responsabilidad. De lo anterior se sigue que la aplicación de este régimen exige, en primer

término, identificar el riesgo de que se trata y, en segundo término, determinar quiénes se encuentran protegidos por él y frente a qué categorías de daño.

Esta delimitación del riesgo guarda estrecha relación con la regulación de la IA. La técnica legislativa empleada tanto en el proyecto de ley chileno como en el derecho comparado ha consistido en clasificar los sistemas de IA según el nivel de riesgo que representan. Esta clasificación resulta particularmente útil para articular un régimen de responsabilidad objetiva, en la medida en que permite la coexistencia de distintos estatutos de responsabilidad, permitiendo que aquellos sistemas que no representan un riesgo anormal podrían regirse por el estatuto común de responsabilidad subjetiva, mientras que aquellos sistemas altamente peligrosos podrían quedar cubiertos por un régimen objetivo (Miranda Valdebenito, 2025, p. 107).

La necesidad de esta distinción se explica por la considerable heterogeneidad que presentan las aplicaciones de IA, y por la consecuente disparidad de los riesgos que cada una de ellas comporta. Así, por ejemplo, no parece justificarse la calificación de riesgo anormal respecto de un asistente de voz, mas sí respecto de un vehículo de conducción autónoma o de sistemas destinados a funciones médicas, en la medida en que estos últimos representan riesgos significativos (Miranda Valdebenito, 2025, p. 107).

Sentada esta clasificación, corresponde ahora precisar el criterio normativo que permite determinar, en cada caso concreto, si un daño específico queda o no comprendido dentro del riesgo cubierto. En este orden de ideas, el fin de protección de la norma opera como un criterio de delimitación conforme al cual la responsabilidad objetiva presupone un riesgo específico, propio de una determinada actividad o cosa, de manera que el daño causado debe corresponder precisamente a ese riesgo. A este respecto, Kótz y Wagner (2006, p. 201) advierten que no corresponde razonar en términos de una probabilidad *ex ante* de ocurrencia del daño, puesto que se trata de una cuestión eminentemente normativa, que opera como un filtro del daño indemnizable equiparable, en su función, al que cumplen los deberes de cuidado dentro de la responsabilidad por culpa.

De este modo, la responsabilidad objetiva cubre únicamente aquellos daños que se insertan dentro de una dinámica de riesgo extraordinario, dinámica que es, precisamente, la que justifica la existencia de un régimen de esta naturaleza. Se configura, así, una esfera de protección circunscrita

a determinadas personas y a determinadas categorías de daño, quedando excluidas de ella las hipótesis ordinarias, que permanecen sujetas al régimen común de responsabilidad. Este criterio constituye, por tanto, el elemento central para la determinación de la atribución de responsabilidad objetiva en el ámbito de la IA (Miranda Valdebenito, 2025, pp. 103-129).

A falta de una norma legal que prevea expresamente un régimen objetivo de responsabilidad, como ocurre en el caso chileno, es preciso analizar el régimen subjetivo. A este respecto, el artículo 2329 del Código Civil establece una presunción de culpa por el hecho propio, norma que ha sido interpretada por la doctrina en relación con dos tipos de situaciones. Por una parte, aquellas relativas a actividades particularmente peligrosas, y por la otra, aquellas en que las circunstancias permiten inferir *prima facie* que el daño ha sido causado por negligencia (Barros Bourie, 2020, pp. 24-31). Considerando esa lectura, la pregunta es si las fases de diseño, creación y desarrollo de sistemas de IA podrían encontrarse dentro de esta hipótesis. Si así fuera, podría configurarse una presunción de culpa respecto de quienes participan en su implementación, quienes solo podrían exonerarse demostrando que el daño provino de factores externos y no de defectos inherentes al diseño o funcionamiento del sistema algorítmico.

Este enfoque permite reconducir el análisis hacia la conducta humana. En lugar de intentar imputar responsabilidad a la máquina, el derecho examina la diligencia de quienes la diseñaron, fabricaron, comercializaron o utilizaron. Concebidas como instrumentos potencialmente riesgosos, estos sistemas imponen a sus responsables la obligación de mantener estándares rigurosos de seguridad, supervisión y monitoreo. De lo contrario, un daño ocasionado por el sistema puede ser interpretado como una manifestación de negligencia en el cumplimiento del deber de cuidado.

Schnake Muñoz (2024, pp. 53-88) sostiene que la decisión, recomendación o predicción que emana de un sistema de IA debe ser siempre atribuible a los actores que tenían el deber de controlarla, lo que lo lleva a proponer la integración de principios internacionales de supervisión humana como fundamento de la responsabilidad. Por su parte, Díaz y Mitrani (2025, pp. 171-184) sugieren un modelo híbrido que combine elementos objetivos y subjetivos según la autonomía del sistema y el grado de control humano efectivo. La responsabilidad objetiva prevalecería en sistemas altamente autónomos, mientras que en sistemas supervisados, se mantendría el filtro de la culpa. Me parece que esta última propuesta tiene una virtud que las posiciones extremas no siempre

ofrecen, reconoce que la IA no es una categoría monolítica, y que la respuesta adecuada depende del grado de control humano que cada sistema permite en cada contexto de uso.

Así las cosas, es preciso mencionar que, la propuesta contenida en la Resolución del Parlamento Europeo de 2017, sugirió la eventual creación de una personalidad electrónica para ciertos sistemas autónomos altamente sofisticados, con el propósito de permitirles responder jurídicamente por los daños que pudieran ocasionar. La iniciativa buscaba ofrecer un marco normativo capaz de gestionar los riesgos derivados del uso de robots inteligentes y sistemas algorítmicos avanzados (Carbajales & Henríquez Herrera, 2024, pp. 1-26). Si un sistema puede tomar decisiones con consecuencias jurídicas, podría parecer lógico dotarlo de un estatus normativo que permita imputarle directamente esas consecuencias.

Suárez-Muñoz (2024, pp., 214-229), ha interpretado esta propuesta como una técnica instrumental de gestión de riesgos, análoga al modelo utilizado para las personas jurídicas, en virtud de la cual, el robot o sistema autónomo funcionaría como un centro de imputación normativa que permitiría canalizar la responsabilidad por los daños que produce. En efecto, el ordenamiento reconoció capacidad de goce y de ejercicio a entidades como las sociedades y corporaciones no porque estas tuvieran, en sentido estricto, voluntad propia, sino por razones eminentemente funcionales.

Esa ficción resultó indispensable para el desarrollo del comercio y de la actividad económica moderna. La personalidad jurídica fue una decisión normativa adoptada para resolver, entre otros, un problema práctico de atribución de responsabilidad. Cabe entonces preguntarse si un razonamiento análogo podría justificar el reconocimiento de una personalidad electrónica a los sistemas de IA, atribuyéndoles capacidad para responder patrimonialmente por los daños que su funcionamiento autónomo pudiera ocasionar.

El centro de decisión que el derecho identifica en su interior está siempre compuesto por personas naturales, capaces de deber, de infringir deberes de cuidado y de ser objeto de reproche jurídico por su propia conducta. La atribución que allí opera no crea una voluntad nueva, sino que imputa a la organización los actos de individuos que ya eran susceptibles de responsabilidad. En los sistemas de IA, por el contrario, no existe dentro de su funcionamiento un centro equivalente de decisión humana que actúe en tiempo real, sino que, sus resultados provienen de procesos de

inferencia estadística, producto de parámetros fijados con anterioridad por sus desarrolladores y de datos cuyo tratamiento escapa, en buena medida, al control directo de cualquier persona en el momento en que el daño se produce.

Pretender construir, respecto de la IA, un centro de imputación que supondría inventar una voluntad donde no existe, en circunstancias de que, en el caso de la persona jurídica, esa voluntad ya estaba presente en sus órganos y solo restaba atribuirla formalmente a la organización. Por ello, considero que el desafío no consiste en trasplantar mecánicamente esta técnica de imputación, sino en reconocer que, tratándose de la IA, el centro de decisión relevante para el derecho de daños sigue estando fuera del sistema, radicado en quienes lo diseñan, entrenan, despliegan y supervisan.

Sin embargo, me parece pertinente recoger la evolución que ha experimentado la discusión en torno a la capacidad de las personas jurídicas para responder extracontractualmente. Como explica Barros Bourie (2020, p. 69), mientras se concibió la culpa como un juicio de reproche estrictamente subjetivo, la responsabilidad de la persona jurídica carecía de fundamento suficiente, particularmente bajo la doctrina de la ficción, conforme a la cual autores como Von Savigny negaban a estas entidades la capacidad de incurrir en delito o cuasidelito, tanto civil como penal. Frente a esa postura, la doctrina de la realidad sostuvo, en cambio, que el órgano no actúa como un tercero ajeno a la persona jurídica, sino que expresa su propia voluntad.

Desde esta perspectiva, la responsabilidad de la persona jurídica por el hecho propio no es sino el resultado de una atribución que el derecho realiza respecto de los actos de sus órganos, agentes y representantes, en la medida en que estos actúan dentro de esos centros de decisión y, por idéntica razón, la persona jurídica responde también por hechos que resultan atribuibles a su organización considerada en abstracto (Barros Bourie, 2020, p. 69).

Por esto, me inclino por sostener que la vía más adecuada no consiste en construir una nueva categoría de sujeto de derecho para la IA, sino en fortalecer los criterios de imputación aplicables a las personas, naturales o jurídicas, que intervienen a lo largo de su cadena de desarrollo y utilización. El problema de la capacidad que plantea el agente algorítmico no se resuelve otorgándole una subjetividad artificial, sino precisando con mayor rigor quién, dentro de esa cadena, debe responder cuando el sistema causa un daño que ningún ser humano decidió

directamente, pero que tampoco habría ocurrido sin la intervención humana previa que hizo posible su funcionamiento.

Finalmente, cabe hacer referencia a otra vía que ha sido explorada en el derecho comparado para articular la atribución objetiva de responsabilidad frente a los daños causados por sistemas de IA; la aplicación analógica de los regímenes especiales de responsabilidad por el hecho de las cosas y, en algunos casos, por el hecho de los animales (Atienza, 2022, pp. 229-254). Respecto de la primera, se ha planteado que ciertos sistemas de IA podrían ser asimilados a la categoría de cosas riesgosas bajo una interpretación extensiva de las normas que regulan la responsabilidad del guardián o del custodio de la cosa que causa daño. Respecto de la segunda, la analogía se funda en que, al igual que un animal, un sistema de IA autónomo puede actuar conforme a impulsos, en este caso algorítmicos, que escapan al control inmediato de su propietario o custodio, lo que justificaría trasladar a este último la responsabilidad por los daños que el sistema ocasione, con independencia de su culpa.

Algunos planteamientos avanzan incluso más allá, sugiriendo que estos supuestos podrían configurar hipótesis de responsabilidad vicaria o por el hecho ajeno, en que el sistema de IA sería tratado como un agente cuya conducta se imputa a quien lo controla, de manera similar a como opera la responsabilidad del principal por el hecho de su dependiente (Atienza, 2022, pp. 229- 254). Sin embargo, esta última construcción presenta una dificultad que ya he tenido oportunidad de examinar. La responsabilidad por el hecho ajeno presupone, precisamente, la existencia de un hecho propio de otro sujeto de derecho, cuya conducta se atribuye a quien debía vigilarlo o dirigirlo. Ello exigiría reconocer a la IA una personalidad jurídica propia, alternativa que, no resulta sostenible.

Por esta razón, considero que estas analogías, aunque valiosas como ejercicio comparativo y como evidencia de la insuficiencia de las categorías tradicionales frente a los riesgos de la IA, no logran sustituir la necesidad de construir un régimen de responsabilidad objetiva fundado directamente en el riesgo creado por la actividad tecnológica, imputado a las personas, naturales o jurídicas, que diseñan, despliegan o controlan los sistemas de IA fuerte. Es en esa dirección, y no en la ficción de atribuir hechos propios a un agente artificial, encuentro la vía más coherente para dar respuesta a los desafíos que la autonomía y la opacidad de la IA plantean al derecho de daños.

Todo lo expuesto me lleva a proponer, como conclusión de este apartado, un modelo híbrido de imputación que, lejos de optar excluyentemente por uno u otro régimen, distribuya la carga de la prueba y el fundamento de la responsabilidad según la intensidad del riesgo que cada sistema de IA representa. En los sistemas de riesgo prohibido y en algunos sistemas de IA fuerte o de alto riesgo me parece justificado reconocer un régimen de responsabilidad objetiva, fundado en el riesgo anormal que estas tecnologías introducen en el tráfico social y en la posición de desventaja probatoria que la víctima enfrenta frente a su complejidad técnica. En cambio, tratándose de sistemas de riesgo permitido, considero que el estatuto común de responsabilidad subjetiva resulta suficiente, sin perjuicio de que dicho estatuto pueda verse reforzado, en supuestos de actividades peligrosas o de negligencia manifiesta, por la presunción del artículo 2329 del Código Civil, debiendo acreditar los presupuestos necesarios para responder extracontractualmente.

Este diseño escalonado permite, a mi juicio, conciliar la necesidad de proteger a la víctima frente a los riesgos más graves con la conveniencia de no imponer cargas indemnizatorias desproporcionadas a quienes despliegan tecnologías de bajo impacto. Con todo, no puedo dejar de advertir que la operatividad de este modelo dependerá, en gran medida, de que la arquitectura de los sistemas de IA evolucione hacia estándares más rigurosos de trazabilidad por capas, que permitan reconstruir con mayor precisión el proceso causal que condujo al daño.

2.4. Régimen de responsabilidad civil en el proyecto de ley que regula la IA en Chile

Habiendo examinado en el capítulo anterior el estado general del proceso regulatorio chileno en materia de IA, corresponde ahora detenerme en el régimen de responsabilidad civil que el propio proyecto de ley contempla para los daños derivados del uso de estos sistemas, el cual, adopta una posición que podría calificarse de conservadora, ya que opta por mantener el régimen tradicional de responsabilidad subjetiva, remitiéndose al estatuto general de responsabilidad por culpa o negligencia consagrado en el Código Civil. En particular, los artículos 28 y 29 del proyecto establecen que la reparación de los daños ocasionados por sistemas de IA se regirá por las reglas

generales de la responsabilidad extracontractual, descartando la instauración de un sistema de responsabilidad objetiva o estricta.

Durand-González y Gerbaud-Kern (2025, pp. 319–336) sostienen que la exigencia de acreditar la negligencia del operador podría convertirse en una verdadera prueba diabólica para las víctimas. La autonomía y complejidad de los sistemas de IA hacen extremadamente difícil reconstruir el trayecto causal que condujo al daño, especialmente cuando ese daño es el resultado de un proceso de aprendizaje automático cuyas decisiones intermedias no quedan registradas de manera comprensible ni accesible. El problema se agrava notablemente cuando intervienen múltiples actores en la cadena tecnológica, ya que la atribución directa de responsabilidad se diluye entre eslabones cuya contribución al daño resulta difícil de individualizar.

En el mismo sentido, Amunátegui Perelló (2020, pp. 76 y ss.) ha advertido, que mantener un modelo estrictamente subjetivo de responsabilidad podría terminar reduciendo los casos en que la víctima logra obtener reparación efectiva. El autor sostiene que las categorías clásicas del Derecho civil podrían resultar insuficientes para abordar los daños complejos derivados de procesos de *machine learning*, en los cuales la conducta dañosa no siempre es plenamente previsible para los operadores humanos y, en ocasiones, ni siquiera para los propios diseñadores del sistema.

Comparto estas preocupaciones, y añadiría que el problema no es solo de atribución, sino también de asimetría. La víctima de un daño algorítmico enfrenta, por un lado, la opacidad técnica del sistema que le causó el perjuicio; por otro, la asimetría informativa y económica respecto del operador tecnológico; y, finalmente, la rigidez de un esquema probatorio diseñado para contextos en los que el daño tiene un origen identificable y la causalidad puede reconstruirse con relativa claridad. Esa triple asimetría configura un escenario en que el acceso efectivo a la tutela judicial se convierte, para muchas víctimas, en una simple promesa.

El artículo 30 del proyecto establece que las acciones civiles derivadas de esta ley se tramitarán conforme al procedimiento sumario, atendida la necesidad de resolver con eficacia y rapidez los conflictos propios de un entorno tecnológico particularmente dinámico. Sin embargo, esta misma celeridad procesal puede convertirse, paradójicamente, en un obstáculo adicional para la víctima que debe acreditar el daño y la relación de causalidad en un litigio de alta complejidad

técnica. La opacidad de los algoritmos, la intervención de múltiples agentes a lo largo de la cadena tecnológica y las dificultades prácticas para acceder, conservar y comprender la prueba digital exigen un estándar probatorio que, en muchos casos, resulta difícil de satisfacer dentro de plazos procesales tan breves.

Esta tensión se agrava si se considera que el sistema de responsabilidad sigue estructurado, en principio, sobre una lógica subjetiva, en la que la carga de la prueba recae sobre quien alega el daño, de acuerdo a las reglas generales. En ese contexto, el procedimiento sumario corre el riesgo de convertirse, en la práctica, en una barrera adicional para la tutela efectiva de la víctima, en lugar de facilitar el acceso a la justicia que busca garantizar.

Con todo, el proyecto contempla un catálogo de acciones civiles que merece atención. El artículo 29 faculta a cualquier persona que sufra un daño por la utilización de un sistema de IA para demandar judicialmente, habilitando tres tipos principales de pretensiones: la cesación de los actos generadores de daño, la indemnización de perjuicios, y la adopción de medidas para evitar la continuación de la infracción ante peligros inminentes.

Se contempla también la posibilidad de exigir la publicación de la sentencia condenatoria a costa del infractor en un diario de circulación nacional, medida orientada a reforzar la transparencia y el control social sobre los operadores tecnológicos responsables. Esta última herramienta, aunque de carácter sancionatorio más que reparatorio, tiene un valor disuasivo que no debe subestimarse en una industria en que la reputación corporativa es un activo de considerable importancia económica.

El régimen de responsabilidad civil se entiende, además, sin perjuicio de otras responsabilidades legales, ya sean administrativas o penales. Los operadores pueden enfrentar simultáneamente multas administrativas de hasta 20.000 UTM en casos de infracciones gravísimas, lo que configura un esquema sancionatorio de múltiples capas que, al menos en teoría, debería generar incentivos suficientes para el cumplimiento normativo.

Queda por resolver si dichos incentivos guardan una relación efectiva de proporcionalidad con las ganancias económicas que los operadores obtienen al desplegar masivamente sistemas de IA, y si los organismos encargados de la fiscalización dispondrán de las capacidades técnicas y los

recursos necesarios para ejercer un control real sobre sistemas cuya complejidad, en no pocos casos, excede las herramientas de supervisión actualmente disponibles.

En lo que respecta al sistema de clasificación de riesgos, Amunátegui et al. (2025, pp. 15-21) reportan que la consulta a expertos evidenció dudas sobre la capacidad real de fiscalización de los sistemas de riesgo inaceptable, la falta de claridad respecto de las normas aplicables a los sistemas de alto riesgo y la insuficiencia de los mecanismos previstos para los incidentes graves. A ello se suma la dificultad de clasificar sistemas multipropósito, cuya aplicación puede variar sustancialmente según el contexto de uso, lo que tensiona la lógica misma de una clasificación de riesgo fijada *ex ante*.

Frente a este diagnóstico, los autores proponen una sustitución del enfoque centrado en la clasificación de riesgos por una regulación orientada directamente a los daños que la IA puede ocasionar, recogiendo la conceptualización propuesta por el Instituto Nacional de Estándares y Tecnología de Estados Unidos, que distingue entre daños a las personas, a las organizaciones y al ecosistema, ajustándola a las categorías de daño ya reconocidas en el derecho nacional (Amunátegui et al., 2025, pp. 15-21). Bajo este modelo, la determinación de la responsabilidad se centraría en el proceso de entrenamiento de los sistemas, exigiendo deberes generales de cuidado, como la información, la trazabilidad razonable y la realización de pruebas proporcionales a la previsibilidad del daño, diferenciados según el rol desempeñado por cada interviniente en la cadena tecnológica.

Es interesante la propuesta de un sistema de puerto seguro o *safe harbor*, fundado en estándares técnicos internacionales como ISO/IEC 42001 o el NIST AI RMF, conforme al cual el cumplimiento certificado de dichos estándares constituiría una presunción legal de debida diligencia (Amunátegui et al., 2025, pp. 15-21). Esta alternativa invierte la lógica de la autorización administrativa previa, propia del modelo europeo, y traslada la determinación de la responsabilidad a los tribunales civiles, sobre la base de la prueba técnica del cumplimiento de dichos estándares.

En ese escenario, concentrar el régimen de responsabilidad en el cumplimiento de deberes previos de diligencia por parte del proveedor puede traducirse en una tutela meramente formal, pues la posibilidad real de exigir dicho cumplimiento, fiscalizarlo y hacer efectiva la

responsabilidad correspondiente se vuelve considerablemente más difícil cuando el sujeto obligado se encuentra fuera del alcance efectivo de nuestros tribunales y órganos de fiscalización.

Por esta razón, sostengo que el desafío para Chile no radica tanto en imponer exigencias de diligencia a quien produce estos sistemas, sino en construir un régimen de responsabilidad civil que atienda a su condición de país consumidor de IA. Esto supone desplazar el centro de gravedad de la regulación hacia quienes efectivamente intervienen dentro del territorio nacional. Sobre estos actores resulta posible construir estándares de diligencia exigibles en la práctica, pues son ellos quienes seleccionan el sistema, lo integran a su operación, lo supervisan y se benefician económicamente de su uso dentro del país.

Las implicancias de no advertir esta diferencia son significativas. Si la atribución de responsabilidad se diseña pensando en un productor nacional inexistente o marginal, el sistema corre el riesgo de generar espacios de irresponsabilidad, en los que la víctima no encuentre un sujeto identificable, solvente y jurídicamente exigible dentro de Chile a quien dirigir su pretensión indemnizatoria. En cambio, si el régimen se construye sobre la base de quien despliega y utiliza el sistema en el territorio nacional, se facilita la identificación del legitimado pasivo, se simplifica la prueba de los presupuestos de la responsabilidad y se otorga a la víctima una vía de reparación efectivamente accesible, sin perjuicio de las acciones de repetición que puedan luego dirigirse contra el proveedor extranjero.

En definitiva, el derecho no debe limitarse a operar como un mecanismo sancionatorio represivo. Debe también asumir un rol preventivo y de gobernanza tecnológica, orientado a asegurar que el diseño, desarrollo y despliegue de sistemas de IA se mantenga subordinado a los principios fundamentales del orden jurídico y a la dignidad de la persona humana. Esa función exige capacidades institucionales que hoy Chile está comenzando a construir, y cuya consolidación requerirá tiempo, recursos y una conversación pública.

De todo lo expuesto, concluyo que la eficacia de la futura ley no dependerá, entonces, de la sofisticación de su catálogo de acciones civiles ni de la severidad de su régimen sancionatorio administrativo, sino de si el legislador logra, antes de su aprobación definitiva, cerrar esa brecha entre la naturaleza del riesgo que pretende regular y los instrumentos jurídicos que pone a disposición de quienes lo sufren.

2.5. Conclusiones preliminares

Al finalizar este capítulo, considero posible concluir lo siguiente:

1. La responsabilidad civil frente a la IA cumple una doble función que no admite jerarquías rígidas. A lo largo de este capítulo constaté que la función preventiva y la función reparatoria de la responsabilidad civil no compiten entre sí, sino que se necesitan mutuamente en el contexto tecnológico. La primera opera como un sistema descentralizado de incentivos que orienta la conducta de desarrolladores y operadores antes de que el daño ocurra; la segunda garantiza que, cuando la prevención falla, exista una respuesta efectiva para la víctima concreta. Ninguna regulación de la IA que privilegie una función en desmedro de la otra estará, a mi juicio, completa.
2. El principio de reparación integral es plenamente aplicable a los daños causados por IA, pero su operatividad depende de superar la opacidad algorítmica. Sostuve que la vigencia de este principio no puede examinarse en abstracto, dado que, la autonomía y la opacidad de estos sistemas condicionan directamente la posibilidad de que la víctima ejerza, en la práctica, su derecho a ser indemnizada. En este punto, identifiqué en la arquitectura de capas desarrollada por Rumelhart una vía concreta, la de la IA explicable, capaz de convertir la trazabilidad algorítmica en un verdadero estándar de diligencia exigible, y no en una mera aspiración declarativa.
3. Ni el régimen subjetivo ni el objetivo, aplicados en su forma pura, resuelven satisfactoriamente el problema de la IA. El análisis de ambos sistemas de atribución me permitió concluir que la responsabilidad objetiva se justifica frente a sistemas de IA fuerte que introduce riesgos anormales, como cuando el agente pierde el dominio efectivo de la decisión dañosa, mientras que la responsabilidad subjetiva, reforzada por la presunción de culpa del artículo 2329 del Código Civil, mantiene utilidad para sistemas de menor intensidad de riesgo. De ello propuse un modelo híbrido y escalonado, que distribuye el

régimen aplicable según la naturaleza y peligrosidad concreta de cada sistema, evitando tanto la desprotección de la víctima como la disuasión injustificada de la innovación.

4. El proyecto de ley chileno, al optar por un régimen íntegramente subjetivo, no se hace cargo de las asimetrías estructurales que la IA introduce. Advertí que la combinación entre responsabilidad subjetiva, procedimiento sumario y opacidad técnica configura una triple asimetría, informativa, económica y probatoria, que puede convertir la tutela judicial en una promesa más formal que real. En efecto, el legislador chileno ha resuelto un problema de diseño regulatorio trasladando, casi en su totalidad, el costo de la incertidumbre tecnológica a la víctima.

5. El desafío chileno es de adecuación, no de imitación. Concluí que la eficacia del futuro régimen de responsabilidad no dependerá de la sofisticación de sus acciones civiles ni de la severidad de sus sanciones administrativas o ministeriales, sino de si el legislador logra construir un sistema pensado desde la posición real de Chile como país consumidor, y no productor, de tecnologías de IA, desplazando el centro de gravedad regulatorio hacia quienes efectivamente despliegan y controlan estos sistemas dentro del territorio nacional.

CAPÍTULO 3: IA Y RESPONSABILIDAD EN LA LABOR JURÍDICA

3.1. Los sistemas expertos o sistemas basados en conocimiento

Los sistemas de IA han sido concebidos para operar con distintos grados de autonomía, utilizando como insumo datos e información que les es proporcionada. Con esos materiales, son capaces de inferir la manera de alcanzar determinados objetivos mediante técnicas de aprendizaje automático o modelos basados en reglas lógicas y estructuras de conocimiento, generando como resultado predicciones, recomendaciones o incluso decisiones orientadas a resolver problemas específicos.

Dentro de esta categoría tecnológica me interesa destacar especialmente los denominados sistemas basados en el conocimiento o sistemas expertos, programas informáticos que emplean una base de conocimiento estructurada para ofrecer métodos de resolución de problemas complejos, diseñados para ejecutar de forma automatizada tareas cognitivas que, tradicionalmente, han sido desempeñadas por especialistas humanos, reproduciendo patrones de razonamiento propios de determinados campos profesionales (Romero & Rodríguez, 2004, pp. 37-44).

En el ámbito jurídico, estos sistemas han encontrado un terreno especialmente fértil. Su capacidad para capturar, organizar y utilizar información proveniente de múltiples fuentes les permite facilitar la resolución de problemas complejos (Romero & Rodríguez, 2004, pp. 37-44), lo que en el derecho se traduce en la posibilidad de procesar extensos repositorios de jurisprudencia, legislación y doctrina, permitiendo al profesional acceder a información a gran escala mientras el sistema interpreta el contexto en que esos datos se insertan. La diferencia con un buscador convencional es que un sistema experto no solo recopila información, sino que la organiza en función de una estructura de conocimiento que le permite razonar sobre ella, al menos dentro de los parámetros para los que fue entrenado.

En la labor profesional, estas herramientas ofrecen ventajas competitivas que sería ingenuo ignorar. Los sistemas basados en conocimiento permiten gestionar grandes volúmenes de información no estructurada, contribuyendo a mejorar la consistencia y eficiencia en la toma de decisiones. Frente a la creciente sobrecarga de trabajo que enfrentan tanto los tribunales como los abogados, la tecnología se presenta como un instrumento de apoyo cuya utilidad es difícil negar. Como señala Atienza (2022), la IA tiende a sustituir principalmente las tareas denominadas "funciones 3D: *dull, dirty y dangerous*", esto es, actividades tediosas, sucias o peligrosas. La categoría de tareas tediosas incluye, por ejemplo, labores administrativas o repetitivas como la redacción de demandas ejecutivas estandarizadas, la tramitación masiva de desarchivos, la recopilación de datos o la revisión preliminar de antecedentes para diseñar una estrategia de litigación. Frente a ese tipo de funciones, no creo que exista objeción en que la IA asista o incluso reemplace al profesional humano, siempre recordando que el sujeto a cargo mantiene un deber de diligencia.

Ahora bien, el riesgo se presenta cuando la comodidad tecnológica se transforma en desidia profesional, cuando el abogado deja de ser el filtro de los resultados que el sistema produce y comienza a delegarle funciones que exigen un juicio de valor, sin la supervisión adecuada. Si bien la IA puede gestionar grandes volúmenes de información con una eficiencia que ningún profesional humano puede igualar, no es capaz de ponderar criterios de justicia y equidad en el caso concreto. La IA procesa datos; el ser humano decide sobre la justicia. Esa distinción es, a mi juicio, el fundamento sobre el cual debe construirse el estándar de diligencia exigible al profesional que utiliza estas herramientas.

Es esa misma distinción la que obliga al derecho de daños a ponerse al día. Considero que se trata de una consecuencia lógica y, en cierto sentido, previsible; ningún sistema de IA es infalible, y su incorporación a la labor jurídica sin los resguardos adecuados hace prácticamente inevitable que, tarde o temprano, se produzcan errores. El problema surge, precisamente, cuando estos sistemas fallan, ya sea por un uso inadecuado por parte de quien los emplea, o por defectos inherentes al propio sistema, sean previsibles o imprevisibles. Si esos errores derivan en un resultado dañoso, deben ser reparados.

Es razonable anticipar, en consecuencia, que el aumento del uso de la IA en la comunidad jurídica trae aparejado un incremento potencial de situaciones de impericia, fallas técnicas y otros

escenarios susceptibles de generar perjuicios, entre los que incluyo afectaciones al derecho a la defensa, vulneraciones a la privacidad, restricciones indebidas a la libertad personal y la pérdida de una oportunidad legítima de obtener justicia. Corresponde, entonces, analizar con mayor detalle cada una de estas situaciones, a fin de determinar cómo el sistema de responsabilidad puede, y debe, hacerse cargo de ellas.

3.2. Responsabilidad de los abogados por los daños derivados del uso negligente de IA en el ejercicio de la profesión

En el caso de los profesionales, el estándar de cuidado exigible no es el de un ciudadano común, es el del especialista que ha asumido voluntariamente una función que otros no pueden desempeñar por sí mismos y que, precisamente por eso, genera en su contraparte una expectativa legítima de competencia y diligencia. Esa expectativa es el núcleo de la responsabilidad civil profesional, definida como "aquella en que pueden incurrir quienes ejercen una determinada profesión al faltar a los concretos deberes que esta impone" (Alterini, et al., 1995, pp. 341-342). Resulta lógico, entonces, que quien se desempeña como profesional deba poseer los conocimientos teóricos y prácticos necesarios para ejecutar su labor con la diligencia y prudencia que corresponde (Honores et al., 2018, pp. 23-68).

Ahora bien, el siguiente análisis excluye aquellos daños que involucren a un sistema de IA por un defecto de fábrica, si sus sesgos fueron incorporados en la fase de entrenamiento por el desarrollador, o si la herramienta funcionó exactamente como fue diseñada, pero ese diseño era intrínsecamente riesgoso. Lo que me interesa es referirme a los abogados que eligen usar una herramienta de IA, que deciden confiar en sus resultados y que afectan los derechos de su cliente y, en algunos casos, la correcta administración de justicia.

El creciente uso de herramientas de IA en el ejercicio profesional del derecho ha comenzado a evidenciar fallas en la práctica judicial que no pueden seguir siendo tratadas como anécdotas. Se está afectando intereses patrimoniales de clientes y, más grave, vulnerando principios estructurales

del proceso, como la veracidad de la información, la lealtad procesal y el correcto funcionamiento de la administración de justicia (Hernández, 2020, pp. 7-27). Frente a eso, resulta imprescindible examinar cómo se configuran estos daños, cuáles son sus consecuencias y de qué manera el derecho de la responsabilidad civil debe responder.

La jurisprudencia comparada es ilustrativa. En Estados Unidos, por ejemplo, en 2024 el Tribunal Federal del Distrito Medio de Florida suspendió a un abogado que había presentado un escrito con precedentes jurisprudenciales inexistentes, producto de alucinaciones de una IA generativa. El comité de Quejas de la Corte fue categórico en señalar que si bien la IA se está convirtiendo en una nueva herramienta para la investigación jurídica, nunca podrá reemplazar la responsabilidad del abogado de cumplir con su deber de diligencia razonable de proporcionar al tribunal precedentes precisos que respalden un argumento legal válido (Fernández, 2024).

En Brasil, el Tribunal Regional del Trabajo de la Octava Región sancionó a dos abogadas por insertar deliberadamente *prompts* maliciosos en presentaciones judiciales, según consta en el expediente TORA-2160-00-01.5520.8.03.01. Esta conducta constituye un caso evidente de fraude procesal; no solo se trata de un uso inadecuado de herramientas, sino de una manipulación calculada mediante instrucciones diseñadas para distorsionar resultados.

Recientemente, la Tercera Sala de la Corte Suprema, en rol 23.322-2025 sancionó a una abogada por falsificar doctrina, atribuyendo posiciones jurídicas inexistentes al profesor Juan Andrés Orrego y Jean Pierre Matus. El fallo señala:

> **"**Que la buena fe es un principio fundamental y transversal en nuestro ordenamiento jurídico que, además, ha recibido una consagración expresa en la ley 20.886 en cuyo artículo 2° letra d) se establece como un deber de las partes en la tramitación de los procesos y como una obligación específica del juez de la causa la de *"prevenir, corregir y sancionar, según corresponda, toda acción u omisión que importe un fraude o abuso procesal, contravención de actos propios o cualquiera otra conducta ilícita, dilatoria o de cualquier otro modo contraria a la buena fe"*.
>
> Agrega que la actuación de la abogada:

> “es demostrativa de falta de profesionalidad y resulta contraria a la rectitud y corrección elemental que deben observar los abogados en su desempeño ante los tribunales de justicia, en términos que deben ser desestimadas sus explicaciones de tratarse de un “error involuntario”. En efecto, en casos como éste la buena fe se traduce en el deber que tiene la parte en orden a verificar la efectividad de la información que aporta al tribunal, lo que no hizo”.

Este caso me permite trazar con nitidez la línea que separa la responsabilidad profesional de la responsabilidad civil por daños derivados del uso de la IA. La primera opera en un plano disciplinario, vinculado al cumplimiento de los deberes de rectitud y buena fe procesal, y se configura al margen de que exista o no un perjuicio patrimonial concreto; la segunda, en cambio, exige un daño efectivo y un nexo causal con la conducta negligente del profesional. El fallo anterior, ilustra bien esta distinción, ya que, la sanción recayó sobre la falta de diligencia de la abogada al no verificar la información aportada al tribunal, sin pronunciarse sobre una eventual responsabilidad civil frente al cliente o a la contraparte afectada. Lo que más me interesa rescatar es que la Corte descartó que el llamado error involuntario exima de responsabilidad, pues la buena fe exige, precisamente, verificar la efectividad de lo que se afirma ante el tribunal. Ese mismo criterio, a mi juicio, debe extenderse al uso de la IA.

En Colombia, la Corte Suprema de Justicia, mediante el Auto AC739-2026, impuso una sanción disciplinaria derivada del uso inadecuado de herramientas de IA generativa por parte de un litigante, marcando un punto de inflexión en la relación entre tecnología y responsabilidad profesional. El caso evidencia uno de los riesgos más insidiosos de los modelos generativos, las denominadas alucinaciones, es decir, la producción de información falsa o inexistente presentada con apariencia de veracidad.

El abogado sancionado estructuró un recurso extraordinario de revisión citando seis sentencias que no correspondían a providencias reales y atribuyendo un contenido erróneo al artículo 250 del Código General del Proceso. La Corte calificó la conducta como una actuación temeraria y una negligencia grave e injustificable, pues el profesional omitió el deber elemental de verificar la autenticidad de las fuentes consultadas. A partir de esta decisión, el sistema judicial colombiano ha comenzado a consolidar un estándar de diligencia reforzada que exige, como mínimo, los deberes fundamentales de supervisión humana real sobre el contenido generado,

verificación rigurosa de las fuentes utilizadas y transparencia respecto del uso de asistencia tecnológica en la elaboración de escritos judiciales.

Pero el uso negligente de la IA en el proceso judicial no se agota en la cita de jurisprudencia y doctrina inexistente. Mucho más grave es la posibilidad de que herramientas de IA sean utilizadas para crear o alterar evidencia procesal. Un abogado que utiliza IA generativa para producir documentos, imágenes, audios o correos electrónicos falsos y los incorpora como prueba en un proceso no solo está incurriendo en negligencia profesional, está cometiendo el delito de falsificación de documentos, con las consecuencias penales que ello conlleva.

La facilidad con que los sistemas generativos actuales pueden producir contenido visual o textual indistinguible de material auténtico, convierte este riesgo en algo real, inmediato y concreto. Y si bien la sanción penal por falsificación opera con independencia de la responsabilidad civil, si el cliente resultó perjudicado por la actuación de su abogado, o la contraparte cuya posición procesal fue afectada por prueba fabricada, e incluso el Estado como garante de la correcta administración de justicia, al tener intereses legítimos de reparación, podrán invocar el derecho de daños, que tendrá que ser capaz de atender.

Estas conductas también se extienden a otros actores del proceso judicial que utilizan IA en el desempeño de sus funciones. El perito que elabora un informe apoyándose en herramientas algorítmicas sin verificar sus resultados, el funcionario judicial que utiliza sistemas de predicción para recomendar medidas cautelares, o el fiscal que emplea IA para analizar evidencia digital son ejemplos de situaciones que plantean las mismas dificultades sobre el estándar de diligencia exigible y la atribución de responsabilidad. Si bien no me ocuparé de ellas en detalle creo que el análisis respecto del abogado ofrece un marco conceptual que, con los ajustes que correspondan a cada función, podría proyectarse sobre esos otros escenarios.

Así las cosas, cabe preguntarse cuál es el estándar de diligencia que le resulta exigible al abogado cuando incorpora herramientas de IA en sus funciones dentro de un proceso judicial. La respuesta debe construirse sobre la base de lo que la doctrina y el ordenamiento jurídico chileno ya establecen respecto de la responsabilidad civil profesional. Miranda (2021) señala que, los abandonos, los errores y los descuidos merecen, además del reproche de conciencia y del juicio desfavorable de terceros, una sanción jurídica. El desconocimiento del saber jurídico o la

negligencia en su aplicación redundan en graves conductas del servicio que el abogado presta. Los profesionales del derecho deben hacerse responsables de los daños causados por su impericia y actuar negligente, derivada de faltas graves a las obligaciones mínimas de diligencia que se deben observar en el ejercicio de la profesión. En ese marco, estimo que la incorporación de IA impone al abogado al menos tres deberes específicos que se articulan sobre sus obligaciones generales.

El primero es el deber de competencia técnica, que no solo implica conocer el derecho aplicable, sino que también, es comprender, al menos en términos suficientes, las capacidades y limitaciones de las herramientas tecnológicas que se utilizan. El abogado no puede emplear sistemas de IA sin contar con la formación mínima que le permita evaluar críticamente sus resultados, en concordancia con la exigencia de no asumir asuntos que excedan sus capacidades profesionales (Honores et al., 2018, pp. 23-68). Esto conecta directamente con lo que el cliente espera al contratar a un abogado, la confianza de que el profesional dispone del conocimiento y la capacidad para resolver el encargo, con la convicción de que recibirá un servicio de calidad a un coste apropiado (Fernández, 2014).

El segundo es el deber de verificación, que implica que toda información generada por sistemas de IA debe ser revisada, contrastada y validada antes de ser utilizada en actuaciones judiciales. Este deber no admite excepciones fundadas en la popularidad o difusión del sistema utilizado. El origen tecnológico de la información no excusa los errores previsibles, y la verificación se configura como un estándar ineludible de la práctica jurídica contemporánea.

El tercero, es el deber de información al cliente. El abogado debe informar de manera clara y suficiente sobre el uso de herramientas de IA en la gestión del encargo profesional, incluyendo sus beneficios, riesgos y posibles implicancias, en resguardo tanto de la buena fe contractual y procesal como de la autonomía del cliente en la toma de decisiones sobre su representación.

El Código de Ética Profesional del Abogado (2011) ofrece un soporte normativo relevante para estos deberes. El artículo 22 dispone que el abogado debe evaluar previamente la naturaleza y complejidad del asunto antes de aceptar el encargo; el artículo 25 prohíbe asumir asuntos que excedan sus conocimientos o capacidades técnicas; y el artículo 31 impone el deber de realizar todas las gestiones útiles destinadas a evitar o mitigar los perjuicios ocasionados al cliente cuando se reconoce una actuación negligente. Más directamente aplicable resulta el artículo 28, que

dispone que el abogado debe adelantarse a reconocer la responsabilidad que le resulte por su negligencia, error inexcusable o dolo, allanándose a indemnizar por los daños y perjuicios ocasionados al cliente. Estas normas son un estándar de conducta que refleja una concepción del ejercicio del derecho fundada en la competencia técnica y el compromiso con la responsabilidad frente a las consecuencias negativas de la propia actuación profesional.

Aunque es necesario reconocer que estas normas éticas resultan exigibles principalmente respecto de abogados colegiados, también lo son respecto de aquellos que no forman parte de dichas organizaciones profesionales (Honores et al., 2018, pp. 23-68), por aplicación del artículo 1546 del Código Civil que permite integrar al contenido del vínculo contractual las reglas del Código de Ética como estándares orientadores del ejercicio de la abogacía, en virtud del principio de buena fe contractual y de la naturaleza de la obligación asumida.

Los daños que pueden ocasionarse al cliente por el uso indebido de la IA, pueden ser tanto patrimoniales como extrapatrimoniales. El daño patrimonial al cliente se produce cuando, como consecuencia de errores en escritos elaborados mediante IA no verificada, se pierde una oportunidad procesal relevante o se dicta una resolución desfavorable. El daño moral, por su parte, emerge de la afectación a la confianza depositada en el abogado, elemento esencial de la relación profesional (Fernández, O., 2014). Asimismo, pueden configurarse supuestos de responsabilidad frente a terceros, cuando la introducción de información falsa en el proceso afecta la posición jurídica de la contraparte o entorpece la labor jurisdiccional. Y, además, no puede descartarse la existencia de daños sistémicos, en la medida en que la proliferación de estas prácticas erosiona la confianza pública en la administración de justicia.

Considero que el marco normativo vigente, fundado principalmente en la responsabilidad contractual derivada del mandato, ofrece herramientas suficientes para abordar estos supuestos, siempre que se interprete de manera evolutiva y acorde con los nuevos contextos tecnológicos (Honores et al., 2018, pp. 23-68). Aunque, es cierto que la complejidad de estos escenarios podría justificar el desarrollo de criterios más específicos, que permitan precisar los estándares de conducta exigibles.

El ejercicio del derecho trasciende el ámbito de una prestación de servicios privados. Se proyecta sobre bienes jurídicos de especial relevancia social, como el acceso a la justicia, la tutela

judicial efectiva y la confianza pública en las instituciones. De ahí que el debate no solo deba centrarse en la suficiencia de las normas actuales, sino también en la eventual necesidad de establecer estándares regulatorios más amplios y uniformes, aplicables a todos quienes ejerzan la profesión jurídica con independencia de su colegiatura, que aseguren que la incorporación de tecnologías avanzadas no comprometa los principios de diligencia, lealtad y responsabilidad que históricamente han estructurado la función social del abogado. La IA puede ser una herramienta extraordinariamente poderosa al servicio de la justicia. Pero solo si quien la opera sabe que, al final del proceso, la responsabilidad sigue siendo suya.

Un antecedente reciente que considero indispensable incorporar a este análisis es la Guía de sugerencias para el uso responsable de IA en el ejercicio profesional, aprobada por el Consejo General del Colegio de Abogados de Chile A.G. en julio de 2026. Se trata, a mi juicio, de la primera respuesta institucional gremial frente a un problema que, como he venido sosteniendo a lo largo de este capítulo, ya se manifestaba en la práctica forense antes de que existiera un marco normativo o deontológico que lo abordara expresamente. El propio documento reconoce esta finalidad orientadora, al señalar que constituye una herramienta de apoyo para el uso responsable de sistemas de IA, sin crear obligaciones adicionales a las ya contempladas en el Código de Ética Profesional (Colegio de Abogados de Chile A.G., 2026, p. 1).

Me parece significativo que la Guía adopte, desde sus principios generales, una posición que coincide plenamente con la distinción que he venido trazando entre la asistencia tecnológica legítima y la delegación indebida del juicio profesional. En efecto, el documento establece que la IA constituye un instrumento al servicio del ejercicio profesional del abogado, que en ningún caso lo sustituye, y que este conserva siempre la responsabilidad personal e indelegable, sin poder excusarse invocando el resultado producido por el sistema (Colegio de Abogados de Chile A.G., 2026, p. 1). Esta afirmación resulta consistente con el estándar de diligencia que he sostenido a lo largo de este capítulo. Sin embargo, no distingue entre la responsabilidad civil y profesional.

El aporte más relevante de la Guía, a mi entender, se encuentra en la sistematización del deber de verificación. El documento precisa que este deber es personal e indelegable, y que no se satisface por el mero funcionamiento del sistema ni por la verificación efectuada por otro sistema de IA, exigiendo al abogado contrastar personalmente toda cita normativa, jurisprudencial, doctrinal o fáctica antes de incorporarla a un escrito o actuación profesional (Colegio de Abogados

de Chile A.G., 2026, p. 2). Esta precisión me parece valiosa porque cierra una vía de escape que, de otro modo, podría invocarse con facilidad, la idea de que basta con que el propio sistema, o una segunda herramienta de IA, revise el resultado producido, para entender cumplido el estándar de diligencia. La Guía descarta expresamente esa posibilidad y reafirma, que la verificación es una tarea humana no delegable.

En la misma línea, el documento incorpora una prohibición explícita de delegar en sistemas de IA las funciones que suponen ejercicio de juicio profesional, entre ellas la formulación de la estrategia de defensa, la ponderación de alternativas de acción y la redacción final de los escritos que el abogado asume como propios (Colegio de Abogados de Chile A.G., 2026, p. 2). Esta disposición traduce en un lenguaje normativo la distinción que propuse anteriormente entre las tareas tediosas, susceptibles de automatización sin mayor reparo, y aquellas que exigen un juicio de valor irreductible a la lógica algorítmica.

Respecto de las actuaciones procesales, el abogado conserva íntegra la responsabilidad por la veracidad de las citas y antecedentes que incorpora a sus escritos, aun cuando estos hayan sido sugeridos o producidos por un sistema de IA, y debe rectificar a la mayor brevedad cualquier información inexistente o inexacta que advierta haber aportado (Colegio de Abogados de Chile A.G., 2026, p. 3). Considero que esta disposición zanja, al menos en el plano deontológico, cualquier intento de invocar el error involuntario como causal de exoneración.

Ahora bien, más allá de estos aportes, la Guía dedica una atención considerable a materias que, si bien relevantes, se alejan del núcleo de la responsabilidad civil que constituye el objeto de esta investigación. Los apartados relativos a la confidencialidad de la información del cliente, al consentimiento informado para el uso de sistemas de IA y a la responsabilidad por proveedores externos configuran, en mi opinión, un estatuto robusto de protección de datos y de gestión de riesgos contractuales (Colegio de Abogados de Chile A.G., 2026, pp. 3-5), pero se sitúan primordialmente en el terreno de la responsabilidad disciplinaria y, eventualmente, contractual, antes que en el de la responsabilidad extracontractual.

La Guía construye, con notable precisión, un estándar de conducta exigible al abogado que utiliza sistemas de IA, pero no se pronuncia sobre las consecuencias civiles que se siguen de su infracción. El texto es enfático en afirmar que el abogado conserva la responsabilidad personal e

indelegable por sus actuaciones (Colegio de Abogados de Chile A.G., 2026, p. 1), pero no distingue, entre la sanción disciplinaria que puede seguirse de la infracción a estos deberes y la eventual obligación de indemnizar los daños que dicha infracción cause al cliente o a terceros. Esta omisión es, a mi juicio, comprensible atendida la naturaleza del instrumento y la entidad que lo emite, pero no deja de evidenciar la importancia de preguntarse: ¿Que ocurre si de un mismo hecho deriva una infracción al deber general de cuidado y en particular, a los deberes profesionales del abogado?

Con todo, no puedo dejar de valorar el impacto que este documento puede tener precisamente para la determinación del estándar de diligencia en sede de responsabilidad civil. Aun cuando la Guía no persigue directamente ese objetivo, su contenido ofrece un catálogo de conductas exigibles, el deber de competencia tecnológica, el deber de verificación personal y la prohibición de delegar el juicio profesional, que bien puede operar como parámetro de referencia para que un tribunal civil determine si un abogado actuó o no con la diligencia debida al utilizar herramientas de IA. En este sentido, considero que la Guía, sin proponérselo expresamente, contribuye a precisar el contenido concreto de la *lex artis* del abogado, cuya infracción, cuando cause un daño efectivo, seguirá debiendo canalizarse a través de las reglas generales de la responsabilidad civil extracontractual, de manera independiente a las consecuencias de la responsabilidad profesional que del mismo hecho ilícito pueden derivar.

3.3. Utilización de la IA en el ejercicio de la jurisdicción

Chile no parte de cero en materia de digitalización judicial. La Ley N.º 20.886, publicada en 2016, que creó la Oficina Judicial Virtual, representó un cambio estructural en la forma de tramitar los procesos. El expediente en papel cedió su lugar al expediente electrónico, las notificaciones se desmaterializaron y la presentación de escritos pasó a realizarse a través de plataformas digitales.

Frente a muchos países de la región, ese fue un avance importante. Pero la tramitación electrónica, con toda su importancia, es apenas el punto de partida; organizar documentos en formato digital no es lo mismo que procesar inteligentemente la información que esos documentos contienen. El salto desde la Oficina Judicial Virtual hacia el uso de herramientas de IA en la función jurisdiccional es enorme, y el camino que nos queda por recorrer en materia de prevención de sesgos, auditoría algorítmica, trazabilidad de decisiones y explicabilidad de sistemas es considerablemente más largo que el ya recorrido.

Hecha esa contextualización, me parece importante comenzar por destacar que incorporar IA en la administración de justicia nunca puede equivaler a sustituir al juez. La incorporación de estas herramientas debe entenderse como el tránsito hacia lo que Molina Sandoval (2022, pp. 226-248) denomina inteligencia aumentada. La tecnología actúa como instrumento auxiliar que permite al magistrado concentrar su labor en el análisis sustancial del conflicto y en la construcción de la decisión jurídica, liberándolo de tareas repetitivas y mecánicas. El juez que utiliza IA para sistematizar jurisprudencia, identificar precedentes relevantes o gestionar volúmenes masivos de información no está delegando su función; está optimizando el tiempo que dedica a ejercerla. Lo que no puede hacer, bajo ninguna circunstancia, es delegar en el sistema la decisión del caso sometido a su conocimiento.

En el ámbito de la administración de justicia, el Banco de Desarrollo de América Latina (CAF, 2022) identifica tres grandes categorías de uso de IA. La primera comprende las tecnologías de semántica y gobernanza de datos, orientadas a la interoperabilidad y organización de la información judicial, como el sistema E-Proc en Brasil o Fiscal Watson en Colombia. La segunda abarca los sistemas de automatización y búsqueda inteligente, destinados a reducir tiempos de tramitación, como Radar en Brasil o SAJ Digital. La tercera corresponde a los sistemas expertos y predictivos, cuyo propósito es aproximarse al razonamiento técnico de un especialista humano, como Prometea en Argentina para la automatización de documentos y asistencia jurídica, y COMPAS en Estados Unidos para predecir riesgos de reincidencia penal. Es precisamente esta última categoría la que concentra los mayores riesgos y la que exige, en consecuencia, los estándares de control más rigurosos.

Uno de los avances más concretos en materia de regulación de la IA en el ámbito judicial comparado se ha producido en Colombia. El Acuerdo PCSJA24-12243 de 2024 del Consejo

Superior de la Judicatura estableció lineamientos para el uso responsable, seguro y ético de herramientas algorítmicas en la Rama Judicial, trazando una distinción de fondo entre funciones auxiliares o administrativas, respecto de las cuales su uso se admite como apoyo, y funciones sustantivas propias de la actividad jurisdiccional, en que se prohíbe su utilización. La valoración de pruebas y la interpretación de normas jurídicas quedan, por tanto, exclusivamente reservadas al juez humano (Consejo Superior de la Judicatura, 2024). Esta delimitación responde al principio de que las decisiones que afectan derechos fundamentales no pueden ser delegadas a sistemas automatizados sin que ello implique una abdicación del control racional que la función jurisdiccional exige. En otros términos, la IA puede asistir al juez, pero no puede reemplazar su juicio.

La resistencia que estos sistemas han encontrado en el Poder Judicial chileno no es, a mi juicio, un fenómeno irracional. Lang (2022, pp. 31-39) identifica como factores estructurales que la explican el temor frente a lo desconocido y la percepción de que la tecnología podría sustituir funciones humanas; el apego a prácticas tradicionales que persiste incluso en contextos de modernización, como se advirtió en la propia transición del expediente papel al expediente digital; y el uso inadecuado de los datos, que incluye tanto la insuficiente sistematización de información como el riesgo de que los datos utilizados para entrenar estos sistemas se encuentren sesgados, comprometiendo la calidad y legitimidad de los resultados. Ese tercer factor es el que más me preocupa, porque opera de manera silenciosa; un sistema que reproduce sesgos lo hace con la misma apariencia de objetividad que uno que funciona correctamente.

Hablar de sesgos en la adopción de IA en la administración de justicia se refiere a la introducción de discriminaciones que no provienen de la malicia de sus diseñadores, sino de los datos con que fueron entrenados. Si la jurisprudencia histórica sobre la que se entrenó un sistema de predicción de reincidencia refleja patrones de criminalización desproporcionada de ciertos grupos sociales o étnicos, el sistema aprenderá esos patrones y los reproducirá con la apariencia de neutralidad estadística.

COMPAS, el sistema utilizado en Estados Unidos para predecir riesgos de reincidencia, ha sido ampliamente documentado como un ejemplo de esta dinámica, ya que puede otorgar puntuaciones de peligrosidad más elevadas a grupos vulnerables de manera injustificada (Meneses Correa et al., 2021, pp. 12-15). El problema no es que el algoritmo mienta; es que dice la verdad

estadística de un pasado injusto y la proyecta hacia el futuro con la autoridad de lo matemáticamente calculado.

La prevención de esos sesgos no puede quedar entregada a la buena voluntad de quienes diseñan o adquieren estos sistemas. Exige mecanismos institucionales concretos, como auditorías algorítmicas previas al despliegue, sistemas de detección incorporados en el proceso de validación, trazabilidad de las decisiones que permita reconstruir el itinerario lógico que condujo a un resultado determinado y, en la medida de lo posible, sistemas de capas que descompongan el proceso decisional del algoritmo en etapas auditables.

Esta última herramienta, que en la literatura técnica se conoce como explicabilidad, no resuelve completamente el problema de la caja negra, pero lo hace más manejable; si no podemos ver el interior del sistema, al menos podemos examinar sus capas y aproximarnos a una comprensión de por qué se produjo un determinado resultado. Ese acercamiento, aunque imperfecto, puede ser suficiente para determinar en qué eslabón del proceso ocurrió el error y, con ello, quién debe responder por él.

Esta exigencia de transparencia, más que un imperativo técnico, es una consecuencia directa de los derechos fundamentales que se deben observar en el proceso judicial. El artículo 19 N.º 3 de la Constitución Política consagra el derecho al debido proceso, y ese derecho se torna ilusorio cuando la decisión judicial se sustenta en sistemas cuyo funcionamiento es desconocido para las partes. Si el juez permite que la decisión se apoye en algoritmos opacos protegidos por el secreto industrial, las partes no pueden impugnar ni controlar racionalmente la base de la decisión, vaciando de contenido el principio de contradicción y el derecho de defensa.

A ello se suma la responsabilidad patrimonial del Estado por error judicial, particularmente en materia penal conforme al artículo 19 N.º 7 letra i) de la Constitución; si una sentencia se funda en sesgos algorítmicos o en información incorrecta no verificada, podrían darse los presupuestos para que opere la reparación de los daños causados por una actuación jurisdiccional errónea. Y los artículos 6 y 7 de la Carta Fundamental, que obligan a todos los órganos del Estado a someter su actuación a la Constitución y la ley, confirman el argumento, pues la función jurisdiccional es indelegable, y la potestad de juzgar pertenece exclusivamente a tribunales establecidos legalmente y ejercida por personas investidas de autoridad.

La jurisprudencia comparada permite ilustrar lo que ocurre cuando esos principios no se respetan. En el Tribunal Superior de Pereira, Colombia, un magistrado consultó a ChatGPT sobre si cierta concentración de alcohol en la sangre de la víctima era suficiente para afectar su capacidad de reacción. Con base en esa respuesta, el tribunal concluyó que el grado de embriaguez no era relevante, revocó una absolución previa y condenó al imputado a treinta y dos meses de prisión.

La determinación de los efectos fisiológicos del alcohol en la capacidad de reacción es una cuestión técnica que requiere prueba pericial especializada, no una respuesta probabilística emitida por un modelo generativo sin valor probatorio autónomo. La sustitución de la pericia forense por una consulta algorítmica no es un accidente tecnológico, sino una manifestación de uso negligente de la IA, que compromete el debido proceso, la contradicción de la prueba y la motivación suficiente de la sentencia. En contraste, tribunales colombianos como el de Medellín han utilizado herramientas como *Copilot* para tareas instrumentales, cumpliendo con el deber de transparencia. La diferencia no está en la herramienta, sino en la función que se le asigna.

La Corte Constitucional colombiana abordó estas tensiones en la Sentencia T-323 de 2024, hito en la progresiva incorporación de la IA en la función jurisdiccional. El caso tuvo origen en una acción de tutela sobre acceso a prestaciones de salud de un menor con trastorno del espectro autista, en cuya resolución el juez de segunda instancia reconoció haber utilizado ChatGPT como apoyo en la formulación de argumentos jurídicos. La Corte determinó que no se configuró una vulneración al debido proceso, en la medida en que la decisión se fundó en razonamientos jurídicos tradicionales y el uso de la IA no implicó una delegación de la función jurisdiccional.

Lo interesante de este fallo es que el tribunal asumió una función prospectiva y estableció criterios orientadores, demostrando que la IA no puede sustituir el juicio prudencial del juez, solo asistirlo, debiendo respetar principios como la transparencia, la verificación de la información, la protección de datos personales y la garantía del debido proceso. No se prohibió el uso de la IA, sino que se estableció un modelo de integración subordinada al control humano, sujeta a estándares reforzados de diligencia.

Frente a este panorama, me parece necesario reflexionar sobre una dimensión que el debate sobre la IA en la justicia suele pasar por alto: la cuestión del criterio. Mucho se habla de la capacidad de los sistemas de IA para procesar información, encontrar patrones y generar

respuestas, pero poco se dice sobre algo que esos sistemas no tienen y que resulta irreemplazable: el criterio, la capacidad de elegir entre respuestas igualmente posibles la que es más justa, más coherente con los valores del ordenamiento, más sensible al contexto humano del conflicto.

El trabajo jurisdiccional del presente y del futuro próximo no estará determinado por quien sepa usar mejor el algoritmo, sino por quien tenga la formación, la experiencia y la sensibilidad jurídica necesaria para elegir la mejor respuesta entre las que el sistema produce. La IA puede generar diez versiones de un considerando; el juez es el único que puede saber cuál de ellas hace justicia. Como señala Nieva Fenoll (2018, pp. 128-136), la máquina no dicta sentencias ni señala condenados; esa responsabilidad sigue teniendo un rostro humano, pues la decisión judicial es, en esencia, un acto de conciencia que no puede reducirse a un cálculo estadístico.

En ese contexto, el problema de la formación judicial adquiere una relevancia que no puede subestimarse. En Chile, el Poder Judicial cuenta con programas de perfeccionamiento, habilitación y formación para sus integrantes, pero esos programas no tienen carácter obligatorio. Esa realidad es preocupante en un escenario en que el uso de herramientas de IA en la función jurisdiccional avanza con independencia de si quienes las utilizan cuentan con la preparación técnica necesaria para hacerlo de manera responsable. Un juez que no comprende cómo funciona el sistema que está consultando no puede supervisarlo críticamente ni puede identificar cuándo sus resultados son el producto de un sesgo o de una alucinación algorítmica.

En consecuencia, la Corte Suprema tiene, a mi juicio, un rol institucional que no puede eludir. El artículo 82 de la Constitución le confiere la superintendencia directiva, correccional y económica sobre todos los tribunales de la nación, con las excepciones que la propia Constitución establece. El ejercicio de esas facultades económicas le permite impartir instrucciones generales sobre el funcionamiento del sistema judicial, y eso incluye la posibilidad de ordenar el uso de herramientas de IA bajo principios y estándares definidos de transparencia, responsabilidad y control humano. No se trataría de prohibir la tecnología ni de imponerla indiscriminadamente, sino de establecer el marco bajo el cual su incorporación en el ejercicio de la jurisdicción resulta jurídicamente legítima. Ese ejercicio de facultades económicas sería, además, la respuesta institucional más coherente con la naturaleza del problema, sin necesidad de esperar a que el legislador regule lo que ya está ocurriendo en las salas de los tribunales.

Este contexto es crucial para comprender el Acta N.º 244-2025 de la Corte Suprema de Chile, publicada el 19 de noviembre de 2025, que aprueba el nuevo Código de Ética Judicial. Este texto dedica el artículo 17 a la responsabilidad en el uso de IA y nuevas tecnologías, estableciendo un marco conceptual y operativo que balancea oportunidad y riesgo de manera sofisticada. La norma reconoce que el empleo de estas herramientas constituye una oportunidad para modernizar y optimizar la gestión de los tribunales, pero advierte que su uso exige la máxima cautela.

La propuesta fundamental del artículo 17 aparece en su segundo párrafo: "la decisión judicial es siempre una tarea humana y su responsabilidad es indelegable". Esta afirmación es, a mi juicio, de naturaleza constitucional. Ningún sistema automatizado puede sustituir el juicio de quien ejerce jurisdicción ni condicionar su independencia e imparcialidad. Este principio de responsabilidad indelegable constituye la línea divisoria entre uso lícito y fraude; las máquinas pueden asistir en tareas instrumentales, pero la responsabilidad por decisiones jurisdiccionales permanece siempre con el juez, de manera que delegar en sistemas automatizados funciones como la valoración de la prueba o la decisión del litigio constituiría una violación de este principio fundamental. La norma exige, además, que la magistratura mantenga un control significativo sobre toda herramienta que utilice, comprendiendo su funcionamiento básico, conociendo qué tipo de datos emplea, identificando qué sesgos pueden estar presentes y estando en condiciones de explicar en lenguaje claro cuál fue el papel de la tecnología en su decisión. Un juez que no pueda cumplir con estos requisitos no debe utilizar la herramienta.

Considero particularmente valiosa la enumeración de riesgos específicos que realiza el artículo 17, tales como, la divulgación indebida de datos personales y sensibles, la filtración de información reservada del expediente, la vulneración del secreto profesional, la delegación de funciones críticas, la reproducción y amplificación de sesgos históricos, la generación de respuestas opacas o inexactas que no puedan ser verificadas, así como la dependencia excesiva de proveedores tecnológicos privados o la exposición a fallas de ciberseguridad.

Este catálogo no se limita a la IA, sino que se proyecta sobre otras tecnologías como la realidad virtual para reconstrucciones o las operaciones telemáticas, exigiendo que todas ellas estén sometidas a salvaguardas de seguridad, confidencialidad y respeto de los derechos fundamentales. El artículo individualiza además la responsabilidad, al establecer que comprender y anticipar estos

riesgos es deber de cada juez, quien deberá ponderar en cada caso la proporcionalidad y necesidad del uso de estas tecnologías, sin poder excusarse alegando culpa de proveedores o de la institución.

Las diez orientaciones prácticas que desarrolla la norma concretizan estos principios abstractos, tales son, el deber de informarse previamente sobre las características de la tecnología; la obligación de capacitarse regularmente; la prohibición de incorporar información judicial en sistemas no autorizados; el deber, que considero central, de verificar siempre de manera independiente las respuestas, citas o referencias provenientes de estos sistemas antes de utilizarlas en resoluciones; el uso subordinado de la traducción y transcripción automáticas; el resguardo del debido proceso en operaciones telemáticas; la obligación de dejar constancia de la función cumplida por la tecnología y de que la decisión descansa exclusivamente en la valoración humana; el deber de documentación y conservación de registros que faciliten el control institucional; y, finalmente, el deber de informar cualquier falla, riesgo o sesgo detectado. En definitiva, el artículo 17 representa un equilibrio sofisticado que reconoce valor instrumental a la tecnología, pero rechaza su uso como fuente de autoridad; permite la asistencia, pero prohíbe la delegación de responsabilidad.

Este marco normativo, sin embargo, exigía traducirse en orientaciones prácticas y operativas que permitieran a la judicatura implementarlo en su trabajo cotidiano. Esa labor fue asumida por la Academia Judicial, que en enero de 2026 publicó la Guía Básica sobre IA Generativa para el ejercicio de Juezas y Jueces, elaborada por Ricardo Lillo, Esteban Ruiz y Cristián Soto (Academia Judicial, 2026). El propio documento reconoce, en su presentación, que el uso de la IA al interior de los tribunales ya se está produciendo de manera asistemática e irregular, tanto de forma correcta como incorrecta, lo que confirma que la práctica antecede a la regulación, y esta última llega precisamente para hacerse cargo de un problema que los propios operadores del sistema ya habían identificado empíricamente (Academia Judicial, 2026, p. 1).

La Guía desarrolla con detalle los principios que el artículo 17 del Código de Ética Judicial enuncia de manera más sintética. En materia de supervisión humana, articula las nociones de *human in the loop/human on the loop*, e identifica con precisión el fenómeno del sesgo de automatización, esto es, la tendencia psicológica a confiar acríticamente en que las soluciones ofrecidas por la tecnología son correctas por el solo hecho de provenir de un sistema aparentemente objetivo (Academia Judicial, 2026, p. 13). En materia de transparencia y explicabilidad, vincula

expresamente estos principios con el problema de la caja negra y con el derecho a un debido proceso, sosteniendo que si un juez o un litigante no pueden comprender el razonamiento detrás de un resultado asistido por IA, se vuelve imposible identificar errores, sesgos o motivos de apelación (Academia Judicial, 2026, p. 15), lo que dialoga directamente con la exigencia de explicabilidad por capas que he sostenido a propósito del legado de Rumelhart.

Ahora bien, es en el apartado dedicado expresamente a la responsabilidad donde la Guía roza, sin desarrollar plenamente, el problema que interesa a esta investigación. El documento afirma que la responsabilidad final de las decisiones asistidas por IA recae siempre en el juez, sin que la función jurisdiccional pueda delegarse a un algoritmo, pero reconoce, de manera breve, que la creciente complejidad de estos sistemas plantea la necesidad de una regulación armonizada que determine la responsabilidad civil por los daños que puedan causar, mencionando la eventual responsabilidad del proveedor ante tecnologías defectuosas (Academia Judicial, 2026, p. 18).

Esta mención revela una tensión que el documento no resuelve; radica en el juez toda la responsabilidad disciplinaria y funcionaria por el resultado final, pero no profundiza en cómo se articula esa responsabilidad con la del proveedor de la tecnología, ni en qué supuestos un error judicial derivado de una herramienta defectuosa podría activar la responsabilidad patrimonial del Estado por error judicial, ya examinada a propósito del artículo 19 N.º 7 letra i) de la Constitución.

Se trata, a mi juicio, de la misma omisión que advertí al examinar la Guía del Colegio de Abogados, ya que, se construye con solvencia un estándar de conducta exigible al operador humano, sin proyectar con igual claridad las consecuencias civiles que se siguen de su infracción. Con todo, ambos instrumentos, el gremial y el institucional, cumplen una función que no debe subestimarse, pues precisan el contenido concreto de la *lex artis* exigible tanto al abogado como al juez que utilizan sistemas de IA, ofreciendo un parámetro objetivo que la jurisprudencia podrá emplear para calificar la diligencia de unos y otros.

Estos criterios orientadores ya han comenzado a implementarse. El primer caso en que se manifiesta su aplicación práctica es la sentencia del Juzgado de Letras del Trabajo de Antofagasta, en causa de tutela laboral RIT T-700-2024, de fecha 22 de mayo de 2026, cuyo considerando vigésimo declara explícitamente que, en cumplimiento del deber de transparencia y en

concordancia con el Acta N.º 244-2025, se dejó constancia de que para labores de redacción se utilizó una herramienta de IA de acceso abierto como apoyo meramente instrumental.

El tribunal precisó que dicho uso se limitó a la preparación de la parte expositiva, el orden de la información y la corrección lingüística, bajo supervisión directa y control total del juez, garantizándose que el análisis de la prueba, la fundamentación jurídica y la decisión final fueron producto exclusivo del razonamiento humano. Esta declaración me parece muy relevante, pues por primera vez un tribunal chileno reconoce públicamente el empleo de IA en apoyo de la redacción de la sentencia de manera explícita y formal, creando un precedente institucional que espero se siga replicando.

Ahora bien, el desafío se vuelve más complejo respecto de los tribunales que operan fuera del Poder Judicial, como el Tribunal Constitucional, los Tribunales Tributarios y Aduaneros, el Tribunal de Defensa de la Libre Competencia, los Tribunales Ambientales o el Tribunal de Contratación Pública. Estos órganos jurisdiccionales no están sometidos a la superintendencia de la Corte Suprema, lo que significa que cualquier estándar que esta defina para el Poder Judicial no les resulta directamente aplicable. La pregunta sobre cómo asegurar que esos tribunales incorporen la IA bajo principios equivalentes de transparencia y control humano es una que el ordenamiento jurídico chileno todavía no responde, y cuya solución probablemente requiera intervención legislativa o, al menos, una coordinación institucional que hoy no existe.

En definitiva, la incorporación de la IA en el ejercicio de la jurisdicción es una realidad que exige ser gestionada con la misma exigencia con que el Poder Judicial gestiona cualquier otro instrumento que incide en el acto de juzgar. La IA puede asistir, organizar y optimizar; puede liberar al magistrado de la arqueología documental para que concentre su energía intelectual en el análisis de fondo y la construcción de la decisión justa (Reusser, 2025). Sin embargo, el timón de la legalidad debe permanecer en manos del juez natural, porque el proceso judicial no es una operación de procesamiento de datos, es la búsqueda institucionalizada de la justicia. Y esa búsqueda, por su propia naturaleza, no puede delegarse en ningún algoritmo.

3.4. Conclusiones preliminares

A lo largo de este capítulo he llegado a las siguientes conclusiones:

1. El examen de los sistemas basados en conocimiento me permitió constatar que su utilidad para tareas tediosas o mecánicas es incuestionable, pero que precisamente esa utilidad genera el riesgo de la sustitución del juicio profesional por la comodidad tecnológica. La distinción entre lo que la IA procesa y lo que el ser humano decide debe orientar la construcción de cualquier estándar de diligencia en esta materia.
2. La responsabilidad civil del abogado por uso negligente de IA encuentra en el ordenamiento vigente herramientas suficientes, aunque incompletas. El análisis de los casos chilenos, estadounidenses, brasileños y colombianos revela que los tribunales, sin necesidad de una regulación especial, ya están sancionando la falta de verificación como una infracción al deber de diligencia profesional. Sin embargo, la jurisprudencia, el Código de Ética Profesional y la Guía del Colegio de Abogados, se concentran en construir el estándar de conducta exigible, sin proyectar con la misma claridad las consecuencias civiles concretas de su infracción. Esta asimetría entre la precisión del deber y la vaguedad de sus consecuencias reparatorias constituye, a mi juicio, uno de los vacíos más urgentes.
3. La distinción entre responsabilidad disciplinaria y responsabilidad civil es clave, debido a que ambos regímenes operan sobre presupuestos distintos y no deben confundirse. La sanción disciplinaria por falta de buena fe procesal no requiere acreditar un daño, mientras que la responsabilidad civil sí lo exige, junto con el nexo causal correspondiente. Considero que esta distinción es indispensable para que las víctimas de un daño algorítmico no confundan la reprobación ética del abogado con su obligación de indemnizar.
4. El Acta N.º 244-2025 y la Guía de la Academia Judicial, se ha logrado anticipar a una regulación legal que aún no existe, estableciendo el principio de responsabilidad indelegable del juez como límite infranqueable frente a la automatización. Valoro especialmente que este marco traduzca el principio en deberes verificables, como la constancia de uso y la verificación independiente.

5. Persiste, sin embargo, un vacío en la articulación entre la responsabilidad del juez y la del proveedor tecnológico. Tanto la normativa chilena como los instrumentos comparados examinados radican en el juez toda la responsabilidad funcionaria, pero no resuelven satisfactoriamente cómo esa responsabilidad se relaciona con la eventual falla del sistema utilizado, ni cómo dicho vacío se conecta con la responsabilidad patrimonial del Estado por error judicial reconocida en el artículo 19 N.º 7 letra i) de la Constitución. Se trata, a mi juicio, del mismo problema de fondo detectado respecto de los abogados. El derecho ha avanzado en fijar estándares de conducta, pero no en precisar sus consecuencias reparatorias cuando esos estándares se incumplen por causas ajenas al operador humano.
6. La fragmentación institucional del Poder Judicial chileno amenaza con generar estándares desiguales de protección en relación con aquellos tribunales que escapan de la superintendencia de la Corte Suprema, lo que implica que los avances alcanzados no se proyectan automáticamente sobre toda la función jurisdiccional del país, lo que en mi opinión constituye un riesgo de desprotección para quienes litigan ante esos otros órganos.

CONCLUSIÓN

Inicié este trabajo con una pregunta que parecía simple: cuando el uso de IA en el ámbito jurídico produce un perjuicio, ¿quién debe responder? Al llegar a este punto, luego de recorrer el origen histórico de la tecnología, su recepción en el Derecho comparado, su proyección sobre los regímenes de atribución de responsabilidad civil y su irrupción concreta en el ejercicio profesional del Derecho y de la función jurisdiccional, puedo afirmar que la pregunta no era simple, pero tampoco es irresoluble. Lo que este recorrido me permitió comprobar es que la dificultad no reside en la falta de respuestas jurídicas, sino en la persistente tentación de buscarlas en el lugar equivocado. A continuación expondré las siguientes reflexiones:

1. Las clasificaciones analizadas, la distinción entre IA débil y fuerte propuesta por Atienza, la diferenciación entre sistemas autónomos y supervisados sostenida por Díaz y Mitrani, y los esquemas de riesgo adoptados por la Unión Europea y el proyecto de ley chileno, son herramientas complementarias que permiten construir un ecosistema de atribución diferenciado. Frente a la disyuntiva doctrinaria entre optar por un régimen único, subjetivo u objetivo, para toda la IA, y optar por un régimen escalonado según el riesgo concreto del sistema, me inclino decididamente por esta segunda corriente. La primera postura, aunque más simple, ignora la heterogeneidad radical que separa a un asistente de voz de un vehículo autónomo o de un sistema de apoyo clínico, y termina o sobreprotegiendo a la víctima frente a riesgos triviales, o desprotegiéndola frente a riesgos catastróficos. El modelo escalonado, en cambio, permite que los sistemas de riesgo prohibido queden sujetos a reglas de inalienabilidad; que los de alto riesgo se rijan por responsabilidad objetiva cuando el riesgo sea anormal, y por la presunción del artículo 2329 del Código Civil cuando, siendo igualmente grave, resulte previsible; y que los de riesgo permitido queden sometidos a las reglas generales de la responsabilidad extracontractual, reforzadas excepcionalmente por esa misma presunción en casos de peligrosidad manifiesta o negligencia evidente.

2. La dificultad que presentan los sistemas multipropósito, incapaces de encasillarse establemente en una categoría de riesgo fijada *ex ante*, evidencia el límite estructural de todo modelo clasificatorio construido sobre la función del sistema antes que sobre el daño

que puede ocasionar. Frente a este problema, la propuesta estadounidense de distinguir según el tipo de perjuicio, a las personas, a las organizaciones o al ecosistema, resulta más sólida que la lógica puramente clasificatoria del Reglamento europeo y del proyecto chileno, en la medida en que desplaza el punto de partida del análisis desde una cualidad abstracta del sistema hacia una consecuencia concreta y verificable. Esta reorientación no sustituye el ecosistema de atribución diferenciada expuesto en la conclusión anterior, sino que lo complementa, resolviendo el problema de clasificación que, por sí solo, no logra superar.

3. Descarto, sin reservas, la posibilidad de atribuir responsabilidad jurídica a la propia IA o de dotarla de una personalidad electrónica. A diferencia de la persona jurídica, cuya responsabilidad se construye atribuyendo formalmente a la organización una voluntad que ya existía en sus órganos, la IA carece de todo centro de decisión humana equivalente en el momento en que el daño se produce. Reconocerle subjetividad no resolvería el problema de la imputación, sino que lo desplazaría hacia una ficción vacía. La determinación del sujeto responsable debe seguir, en cambio, un criterio funcional, que distinga entre proveedor, distribuidor, implementador y usuario según a quién correspondía el deber de diligencia incumplido. Cuando el daño provenga de una falla del propio sistema, corresponde orientar la imputación hacia el creador o entrenador del *software*, en un régimen próximo a la responsabilidad objetiva; cuando derive del mal uso del sistema por parte de quien lo opera, corresponde aplicar el régimen subjetivo general, reforzado por la presunción del artículo 2329 en los supuestos de mayor peligrosidad.

4. El problema de la opacidad algorítmica, expresado en la metáfora de la caja negra, no constituye un obstáculo insuperable para la reconstrucción del nexo causal, sino, en buena medida, una decisión de diseño técnico susceptible de exigencia jurídica. La arquitectura de capas ocultas o intermedias que la propia técnica computacional ha desarrollado permite, al menos parcialmente, descomponer la decisión final del sistema en etapas identificables, y con ello determinar en qué fase del proceso se originó o amplificó el error que causó el daño. Esta posibilidad es la que, a mi juicio, permite sostener con mayor solidez la corriente doctrinaria que exige explicabilidad como estándar de diligencia, por sobre aquella que

resigna de antemano toda pretensión de trazabilidad y traslada íntegramente el riesgo probatorio a la víctima.

5. La aplicación analógica de las presunciones de culpa por el hecho de las cosas, particularmente bajo la calificación de cosas peligrosas, constituye una vía doctrinaria valiosa como evidencia de la insuficiencia de las categorías clásicas, pero no la considero plenamente viable como solución autónoma, en la medida en que no logra capturar con precisión la especificidad del riesgo tecnológico ni sustituye la necesidad de un régimen de responsabilidad objetiva construido directamente sobre el riesgo creado por la actividad.

6. La determinación del ámbito jurisdiccional aplicable constituye una cuestión que no puede desatenderse en ningún régimen de responsabilidad que pretenda ser efectivo, particularmente en países que, como Chile, no participan como productores sino como consumidores de estas tecnologías.

7. En el ejercicio profesional del abogado, la distinción entre responsabilidad civil y responsabilidad disciplinaria resulta indispensable. Los juristas que utilizan sistemas de IA deben adoptar las precauciones necesarias para no incurrir en errores o negligencias, cumpliendo deberes de competencia técnica y de verificación personal e indelegable, deberes que se proyectan con igual fuerza sobre los jueces y demás operadores del sistema de justicia, y cuyo incumplimiento debe canalizarse a través de las reglas generales de la responsabilidad civil extracontractual cuando de él se siga un daño efectivo, sin perjuicio de la reprobación ética que pueda corresponder de manera independiente. A ello se suma la necesidad de precaver, desde ya, el riesgo creciente de falsificación o alteración de pruebas mediante estos sistemas, frente al cual tanto abogados como jueces deben responder con estándares reforzados.

8. En definitiva, el análisis desarrollado confirma que las categorías tradicionales del derecho de daños no resultan inservibles, pero exigen una reconstrucción cuidadosa que combine regímenes de atribución diferenciados según el riesgo y el tipo de daño, sin ceder jamás a la tentación de diluir la responsabilidad humana detrás de la complejidad del instrumento. La proyección futura de esta materia dependerá, en último término, de que el legislador logre transformar estos criterios dogmáticos en normas suficientemente flexibles para

adaptarse a una tecnología cuya única certeza es que seguirá cambiando más rápido que el derecho llamado a regularla.

BIBLIOGRAFÍA

Academia Judicial. (2026). *Guía básica sobre inteligencia artificial generativa para el ejercicio de juezas y jueces*. https://www.academiajudicial.cl

Allen, Ronald (1997). Rationality, Algorithms and Juridical Proof: A Preliminary Inquiry, en *The International Journal of Evidence & Proof Northwestern University School of Law* (special). pp. 254–275. https://doi.org/10.1177/1365712797001special01

Alterini, A., Almeal, O., López, R. (1995). *Derecho de Obligaciones Civiles y Comerciales*, Primera edición reimpresa. Buenos Aires, Ed. Abeledo-Perrot. pp., 341-342.

Amunátegui Perelló, C. (2020). *Arcana technicae: El derecho y la inteligencia artificial*. Tirant lo Blanch.

Amunátegui, C., Aránguiz, M., & Jara, A. (2025, diciembre). *Ley de inteligencia artificial: Tensiones y limitaciones en su diseño normativo* (Observatorio Legislativo N.° 44). Centro de Políticas Públicas, Pontificia Universidad Católica de Chile.

Amunátegui Perelló, C., Madrid Ramírez, R., & Aranguiz Villagrán, M. (2021). *Inteligencia artificial y Poder Judicial: Chile y sus desafíos pendientes*. Pontificia Universidad Católica de Chile. Proyecto Preparación del sector judicial para la inteligencia artificial en América Latina, Centro de Estudios en Tecnología y Sociedad (CETyS), Universidad de San Andrés. https://repositorio.udesa.edu.ar/jspui/handle/10908/18634

Araya Paz, C. (2020). Desafíos legales de la inteligencia artificial en Chile, en *Revista Chilena de Derecho.* pp., 257–290.

Ashley, Kevin (2017): Artificial Intelligence and Legal Analytics: New Tools for Law Practice in *The Digital Age*, (Cambridge University Press, Cambridge).

Ataz, J. (2020). *Daños causados por las cosas: una nueva visión a raíz de la robótica y de la inteligencia artificial*, working paper 4/2020, Universidad de Barcelona.

Atienza, M. L. (2022). *Daños causados por la inteligencia artificial y responsabilidad civil*. Editorial Atelier.

Ayo Ferrándiz, Cristina; Seijo Bar, Álvaro; Garre Anguera de Sojo, Ignacio y González Guillén, Pablo (2025). Responsabilidad civil e inteligencia artificial, en *Actualidad Jurídica Uría Menéndez*, 67, pp. 29-56.

Ballesteros, Soledad (1992). La representación del conocimiento en los sistemas conexionistas. En *Psicothema*, (volumen 4, número 2). pp. 343-351. Disponible en_ https://www.redalyc.org/pdf/727/72704202.pdf

Barros Bourie, E. (2020). *Tratado de responsabilidad extracontractual* (2ª ed.). Editorial Jurídica de Chile.

Belda, Ignasi (2019). *Inteligencia artificial: De los circuitos a las máquinas pensantes*. RBA Libros, Barcelona. pp., 80-85.

Bermejo i Alegret. (2026). *Gobernanza de la IA en la UE y España: hitos y retos para 2026*. https://www.bermejoialegret.com/gobernanza-de-la-ia-en-la-ue-y-espana-hitos-y-retos-para-2026/

Brüggemeier, Gert, Haftungsrecht, Springer, Berlin-Heidelberg, 2006.

Caicedo, E., & López, J. (2009). *Una aproximación práctica a las redes neuronales artificiales*. Editorial Univalle. pp., 70-78.

Carbajales, M., & Henríquez Herrera, I. (2024). Personalidad electrónica de "robots inteligentes" como técnica para la reparación de daños: Análisis crítico de las normas de derecho civil sobre robótica del Parlamento Europeo, en *Prudentia Iuris*, (98), pp., 1-26. https://doi.org/10.46553/prudentia.98.2024.3

Carrasco, C. (2026, 3 de febrero). Proyecto de ley de inteligencia artificial en Chile: Qué regula, en qué estado está y por qué importa a las organizaciones, en *Neuronet.*

Campusano Droguett, R. F., & Giesen Espejo, J. B. (2024). La Ley de Inteligencia Artificial de la UE: Un hito normativo en la escena global, en *Actualidad Jurídica* (50), pp., 105–136.

Canaris, Claude y Larenz, Karl, Lehrbuch des Schuldrechts, Beck, München, (1994).

Calabresi, G y Melamed, D. (1972). Property Rules, Liability Rules, and Inalienability: One View of the Cathedral, en *Harvard Law Review*, 6 (85) pp. 1089-1128. Traducido al castellano por el Centro de Estudios Públicos. Recuperado de https://www.cepchile.cl/cep/site/docs/20160303/20160303183635/rev63_calabresi.pdf

Colegio de Abogados de Chile. (2011). *Código de ética profesional*.

Colegio de Abogados de Chile A.G. (2026, julio). *Guía de sugerencias para el uso responsable de inteligencia artificial en el ejercicio profesional*. Colegio de Abogados de Chile A.G.

Comisión de Futuro, Ciencias, Tecnología, Conocimiento e Innovación. (2025). *Segundo informe de la Comisión de Futuro, Ciencias, Tecnología, Conocimiento e Innovación acerca del*

proyecto de ley que regula los sistemas de inteligencia artificial (Boletines refundidos Nos 16821-19 y 15869-19). Cámara de Diputadas y Diputados de Chile.

Consejo Superior de la Judicatura de Colombia. (2024). *Acuerdo PCSJA24-12243 del 16 de diciembre de 2024: Por el cual se adoptan lineamientos para el uso y aprovechamiento respetuoso, responsable, seguro y ético de la inteligencia artificial en la Rama Judicial*. https://actosadministrativos.ramajudicial.gov.co/GetFile.ashx?url=~%2FApp_Data%2FUpload%2FPCSJA24-12243.pdf

Congreso de la República de Colombia. (2025). *Proyecto de Ley Nº. 442 de 2025 Senado: Por medio del cual se regula la inteligencia artificial en Colombia para garantizar su desarrollo ético y responsable y se dictan otras disposiciones* (radicado 7 de mayo de 2025). Publicado en *Gaceta del Congreso No. 0990 de 2025*. https://sidn.ramajudicial.gov.co/SIDN/PUBLICACIONES%20PERIODICAS/TEXTO%20COMPLETO%20Y%20TABLAS%20DE%20CONTENIDO/GACETA%20DEL%20CONGRESO/GC%200990%20DE%202025.PDF

Congreso Nacional de Chile. (2024). *Proyecto de ley que regula los usos de la inteligencia artificial* (Boletín N.º 16821-19). Ingresado en mayo de 2024; en segundo trámite constitucional ante el Senado de la República (2026).

Corte Suprema de Chile. (2025, 19 de noviembre). Acta Nº 244-2025. Código de Ética Judicial. *Diario Oficial de la República de Chile*.

Díaz, M. P., & Mitrani, C. (2025). La responsabilidad por daños causados por el uso de IA: ¿Responsabilidad objetiva, subjetiva o híbrida? en *Revista Jurídica de la Universidad de San Andrés*, (20), pp., 171-184.

Durand-González, J. B., & Gerbaud-Kern, L. (2025). Problemas para la puesta en práctica del régimen de responsabilidad civil en el proyecto de ley de inteligencia artificial (Mensaje N.° 063-371), en *Revista Chilena de Derecho Privado*, 44, pp., 319–336.

Estados Unidos. Presidente. (2023, 30 de octubre). *Executive Order 14110: Safe, Secure, and Trustworthy Development and Use of Artificial Intelligence*. *Federal Register, 88*(210), 75191–75226. https://www.federalregister.gov/documents/2023/11/01/2023-24283/safe-secure-and-trustworthy-development-and-use-of-artificial-intelligence

Fernández, C. B. (2024, 4 de abril). Un tribunal federal norteamericano suspende de ejercicio a un abogado por presentar un escrito procesal elaborado por IA que incluía precedentes

inexistentes en *World Compliance Association*. https://www.worldcomplianceassociation.com/4524/noticia-un-tribunal-federal-norteamericano-suspende-de-ejercicio-a-un-abogado-por-presentar-un-escrito-procesal-elaborado-por-ia-que-incluia-precedentes-inexistentes.html

Fernández, O. (2014). La confianza, fundamento de la relación entre abogado y cliente, en *Legal Today*. https://www.legaltoday.com/gestion-del-despacho/estrategia/la-confianza-fundamento-de-la-relacion-entre-abogado-cliente-2014-02-10/

García, J. M., & Camacho, L. (2024, 13 de diciembre). Regulación de la IA en Brasil: Media sanción para un proyecto de ley que seguirá en discusión, en *Derechos Digitales*. https://www.derechosdigitales.org

Garrigues. (2025, 31 de octubre). *Mapa regulatorio de la inteligencia artificial en Colombia*. https://www.garrigues.com/es_ES/noticia/mapa-regulatorio-inteligencia-artificial-colombia

Hernández, J. (2020). Can machines think? Inteligencia artificial y derecho de daños, en *Revista E-Mercatoria*, 19(1). pp., 7-27.

Honores, M., Vargas, L., & Rojas, J. (2018). *Responsabilidad del abogado en el ejercicio de su profesión* (1.ª ed.). El Jurista. pp., 23-68.

Kótz, Heiny Wagner, Gerhard (2006). Deliktsrecht. München: Luchterhand.

Lang, M. (2022). La Inteligencia Artificial en la Administración de Justicia, en *Ars Iuris Salmanticensis,* vol. 10, Diciembre 2022, pp., 31-39. Ediciones Universidad de Salamanca.

Leavitt, David (2006): *El hombre que sabía demasiado: Alan Turing y la invención de la computadora* (traducción: Federico Corriente). Editorial Antoni Bosch, Barcelona. pp., 181-216.

Manheim, K., & Kaplan, L. (2019). Artificial intelligence: Risks to privacy and democracy, en *Yale Journal of Law & Technology*, 21(1). pp., 111-113.

Mancilla, A., Ebratt, R., & Capacho, J. (2016). *Diseño y construcción de algoritmos*. Universidad del Norte. pp., 1-8.

Meneses Correa, J. S., Pérez Contreras, S. P., & Vaca Salamanca, M. A. (2021). *El debido proceso como límite de la inteligencia artificial*. Universidad Santo Tomás. pp., 12-15.

Ministerio de Ciencia, Tecnología e Innovación, & Ministerio de Tecnologías de la Información y las Comunicaciones. (2025, 7 de mayo). *Proyecto de ley "Por medio del cual se regula la*

inteligencia artificial en Colombia para garantizar su desarrollo ético y responsable y se dictan otras disposiciones". Gobierno de Colombia.

Ministerio de Ciencia, Tecnología, Conocimiento e Innovación. (2024). *Política Nacional de Inteligencia Artificial*. Gobierno de Chile.

Miranda, F. (2021). Responsabilidad civil del abogado, en *Revista del Abogado*, (21).

Miranda Valdebenito, J. (2025). Distribución de riesgos artificiales: hacia una responsabilidad objetiva por daños causados por inteligencia artificial, en *Debates Jurídicos y Sociales*. (11). Disponible en: https://debatesjuridicosysociales.cl/ojs/index.php/djs/article/view/194

Molina Sandoval, S. (2022). La inteligencia artificial en la ejecución de la sentencia civil, en *Revista de Derecho Procesal y de Derechos Humanos*, 1(1), pp., 226–248.

Navas, S. (2017). *Derecho e inteligencia artificial desde el diseño: Aproximaciones,* en Inteligencia artificial, tecnología y derecho. Tirant Lo Blanch.

Nancholas, B. (2023, 1 de septiembre). *Narrow artificial intelligence: Advantages, disadvantages, and the future of AI.* University of Wolverhampton. Disponible en: https://online.wlv.ac.uk/blog/narrow-artificial-intelligence-advantages-disadvantages-and-the-future-of-ai/

Nieva Fenoll, J. (2018). *Inteligencia artificial y proceso judicial*. Marcial Pons. pp., 128-136.

Orrego Acuña, J. A. (2026). *Responsabilidad extracontractual*. https://www.juanandresorrego.cl/apuntes_all.html

Parlamento Europeo y Consejo de la Unión Europea. (2024). *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 (Ley de Inteligencia Artificial)*. Diario Oficial de la Unión Europea. https://eur-lex.europa.eu/eli/reg/2024/1689/oj

Prosser, W., Keeton R., Owen D., Dobbs, Dan y Page, Keeton (1984). *Prosser and Keeton on the Law of Torts*. St. Paul MN: West Publishing, 5a edición bajo la edición general de W. Page Keeton.

Quintana, Fernando (2011). *Interpretación y argumentación jurídica*. Editorial Jurídica de Chile. pp., 304-309.

República de Chile. (1855). *Código Civil de la República de Chile*. Diario Oficial de la República de Chile.

Reusser, C. (2025, 27 de diciembre). Jueces e inteligencia artificial en la justicia civil, en *DerechoInformatico.cl*. https://www.derechoinformatico.cl/ia-en-la-justicia-civil-de-chile/

Rodríguez, P. (2018). *Inteligencia artificial: Cómo cambiará el mundo (y tu vida)* (Trad. J. Rizzo). Deusto. pp., 25-47.

Romero, M., & Rodríguez, J. (2004). Sistemas basados en el conocimiento, en *Revista Vínculos*, 1(1), pp., 37–44. https://doi.org/10.14483/2322939X.4070

Scherer, M. U. (2016). Regulating artificial intelligence systems: Risks, challenges, competencies, and strategies, en *Harvard Journal of Law & Technology*, 29(2). pp., 353–400.

Schnake Muñoz, V. (2024). *Inteligencia artificial y responsabilidad extracontractual: Desafíos frente a la normativa chilena actual*. Memoria de prueba para optar al grado de Licenciada en Ciencias Jurídicas y Sociales, Universidad de Chile.

Suárez-Muñoz, F. (2024). Inteligencia artificial: Una mirada crítica al concepto de personalidad electrónica de los robots, en *División de Ciencias Sociales y Humanidades*, Universidad de Guanajuato. https://doi.org/10.24215/15146774e060

Suñé Llinás, E. (2020). *Derecho e inteligencia artificial: De la robótica a lo posthumano*. Tirant Lo Blanch. pp., 42-44.

Tellez Tejada, N. (2025, febrero). Regulación: Chile introdujo cambios a su política de inteligencia artificial, en *TeleSemana.com*. Disponible en: https://www.telesemana.com/blog/2025/02/06/regulacion-chile-introdujo-cambios-a-politica-nacional-de-inteligencia-artificial/

The White House. (2025, July). *America's AI action plan*. https://www.whitehouse.gov/wp-content/uploads/2025/07/Americas-AI-Action-Plan.pdf

The New York Times. (29 de octubre de 2025). *Character.AI to Bar Children Under 18 From Using Its Chatbots* by Natallie Rocha and Kashmir Hill. Disponible en: https://www.nytimes.com/2025/10/29/technology/characterai-underage-users.html

Torres Mac-Pherson, F., Salgado Alonso, J., & Villarroel Contreras, M. (2025). Responsabilidad civil por daños provenientes de ilícitos derivados del uso de la inteligencia artificial: Régimen aplicable y desafíos de seguridad para su desarrollo, en *Revista Chilena de Derecho y Tecnología*, *14*, e74613. pp., 26-37. https://doi.org/10.5354/0719-2584.2025.74613

Valbuena, R. (2021). *Inteligencia artificial: Investigación científica avanzada centrada en datos*. Editorial: Central Press. ISBN: 978-980-18-1529-7

JURISPRUDENCIA CITADA

Corte Constitucional de Colombia. (2024). *Sentencia T-323 de 2024*. Disponible en: https://www.corteconstitucional.gov.co/relatoria/2024/t-323-24.htm

Corte Suprema de Justicia, Colombia, con fecha 13 de febrero de 2026, en Auto AC739-2026. Disponible en: https://incp.org.co/wp-content/uploads/2026/02/Auto-AC739-2026-del-13-de-febrero-de-2026.pdf

Corte Suprema de Chile, Tercera Sala. Rol 23.322-2025, de 22 de abril de 2026. Disponible en: https://www.pjud.cl/prensa-y-comunicaciones/noticias-del-poder-judicial/145652

Juzgado de Letras del Trabajo de Antofagasta, en causa de tutela laboral RIT T-700-2024, de fecha 22 de mayo de 2026. Disponible en: https://www.diarioconstitucional.cl/wp-content/uploads/2026/05/sentencia_censurada_total_redacted.pdf

Tribunal Superior del Distrito Judicial de Pereira, Sala Penal. (3 de abril de 2024). Sentencia de segunda instancia, Radicado Nº. 76001600019320138073401 [MP: Yarzagaray, M.]. Disponible en: https://www.ambitojuridico.com/sites/default/files/2024-04/Sent-76001600019320138073401-24.pdf

Tribunal Regional del Trabajo de Parauapebas, Brasil. Expediente TORA-2160-00-01.5520.8.03.01. Disponible en: https://www.diariojudicial.com/uploads/0000061635-original.pdf

Tribunal de Distrito de los Estados Unidos, MD Florida. Sentencia de 21 de mayo de 2025 en Caso Megan Garcia v. Character Technologies Inc., Google LLC, Alphabet Inc. N.º 6:24-cv-1903-ACC-UAM. Disponible en: https://www.courtlistener.com/docket/69300919/garcia-v-character-technologies-inc/