



FACULTAD DE DERECHO

**LA RESPONSABILIDAD CIVIL DERIVADA DE LOS DAÑOS OCASIONADOS POR
EL USO DE INTELIGENCIA ARTIFICIAL**

Marta Rodríguez Mora
4º curso grado en Derecho (E-1)
Derecho civil

Madrid
Junio 2026

ÍNDICE

1. INTRODUCCIÓN.....	6
2. LA INTELIGENCIA ARTIFICIAL Y SU RELEVANCIA JURÍDICA.....	8
2.1. Concepto y características de la IA.....	8
2.2. Clasificación de los sistemas de IA según el nivel de riesgo.....	10
2.3. Riesgos jurídicos derivados de su uso: el efecto caja negra.....	13
3. LOS FUNDAMENTOS DE LA RESPONSABILIDAD CIVIL EN EL DERECHO ESPAÑOL.....	16
3.1. Concepto y función.....	16
3.2. Responsabilidad contractual y extracontractual.....	17
3.3. Elementos de la responsabilidad civil extracontractual.....	18
3.3.1. La acción u omisión.....	19
3.3.2. El daño.....	19
3.3.3. La relación de causalidad.....	20
3.3.4. La culpa o negligencia.....	21
3.4. Las tensiones entre el modelo clásico y la IA.....	23
4. EL DAÑO CAUSADO POR SISTEMAS DE INTELIGENCIA ARTIFICIAL.....	24
4.1. Tipología de daños indemnizables.....	24
4.2. Particularidades de los daños causados por sistemas autónomos.....	27
4.3. Problemas de imputación objetiva y subjetiva.....	30
4.3.1. La quiebra del estándar clásico de diligencia.....	30
4.3.2. Los dos escenarios de imputación causal y la insuficiencia de la teoría de la causalidad adecuada.....	31
4.3.3 Criterios causales complementarios.....	32
4.3.4. El sistema de imputación diferenciada según el riesgo.....	33
5. MODELOS DE RESPONSABILIDADES APLICABLES.....	34
5.1. Las opciones regulatorias y la evolución del debate europeo.....	34
5.2. La responsabilidad por culpa y sus limitaciones.....	35
5.3. La responsabilidad por productos defectuosos.....	35
5.4. La responsabilidad objetiva en actividades de riesgo.....	36
5.5. El seguro de responsabilidad civil como mecanismo complementario.....	37
5.6. Propuestas doctrinales de nuevos modelos.....	37

6. EL RÉGIMEN PROBATORIO: LA CARGA DE LA PRUEBA Y SUS PARTICULARIDADES.....	38
6.1. El efecto caja negra como obstáculo probatorio.....	38
6.2. Los mecanismos de exhibición de pruebas como instrumento de equilibrio.....	39
6.3. Las presunciones iuris tantum en la nueva Directiva sobre productos defectuosos....	40
6.4. El régimen probatorio del AI Act y su interacción con la responsabilidad civil.....	41
7. TENDENCIAS REGULATORIAS EN LA UE.....	42
7.1. El marco regulatorio europeo sobre IA: estructura y lógica de conjunto.....	42
7.2. Reglamento europeo de Inteligencia Artificial (AI Act).....	43
7.3. La nueva Directiva sobre responsabilidad por productos defectuosos.....	44
7.4. La Propuesta de Directiva sobre responsabilidad en materia de IA: estado actual y perspectivas.....	44
7.5. La adaptación del Derecho de Daños al contexto tecnológico: valoración crítica.....	45
7.6. Propuestas de lege ferenda: hacia un sistema integrado de prevención, identificación y reparación.....	46
A) Primera propuesta: la certificación de aptitud funcional como requisito habilitante para el despliegue de sistemas de alto riesgo.....	47
B) Segunda propuesta: la identificación obligatoria del usuario como condición de acceso a sistemas de IA generativa.....	49
C) Tercera propuesta: la limitación de responsabilidad diferenciada según el rol en la cadena de valor y la trazabilidad jurídica obligatoria.....	50
D) Consideración de conjunto: las tres propuestas como sistema integral.....	51
8. CONCLUSIONES.....	52
BIBLIOGRAFÍA.....	55
1. LEGISLACIÓN.....	55
2. JURISPRUDENCIA.....	56
3. OBRAS DOCTRINALES.....	56
4. RECURSOS DE INTERNET.....	60

LISTADO DE ABREVIATURAS

AGI — Inteligencia Artificial General (Artificial General Intelligence)

AI Act — Reglamento (UE) 2024/1689, por el que se establecen normas armonizadas en materia de inteligencia artificial

art. / arts. — artículo / artículos

BOE — Boletín Oficial del Estado

CC — Código Civil

CE — Constitución Española

CEE — Comunidad Económica Europea

COM — Documento de la Comisión Europea

coord. / coords. — coordinador / coordinadores

DOCE — Diario Oficial de las Comunidades Europeas

DOUE — Diario Oficial de la Unión Europea

et al. — et alii (y otros)

IA — Inteligencia Artificial

LEC — Ley de Enjuiciamiento Civil

LOPDGDD — Ley Orgánica 3/2018, de 5 de diciembre, de Protección de Datos Personales y Garantía de los Derechos Digitales

núm. — número

op. cit. — opere citato (obra citada)

p. / pp. — página / páginas

RGPD — Reglamento (UE) 2016/679, General de Protección de Datos

STJUE — Sentencia del Tribunal de Justicia de la Unión Europea

TJUE — Tribunal de Justicia de la Unión Europea

TRLGDCU — Texto Refundido de la Ley General para la Defensa de los Consumidores y Usuarios

TS — Tribunal Supremo

UE — Unión Europea

vol. — volumen

1. INTRODUCCIÓN

La inteligencia artificial ha dejado de ser una promesa tecnológica para convertirse en una realidad operativa que atraviesa los ámbitos más sensibles de nuestra vida jurídica y social. Los sistemas autónomos toman hoy decisiones que afectan a diversas áreas, generando consecuencias jurídicamente relevantes para personas concretas sin que, en muchos casos, sea posible identificar con claridad quién ha tomado esa decisión ni reconstruir el proceso que condujo a ella. Esta realidad plantea al Derecho de Daños un desafío que el presente trabajo se propone analizar.

La elección de este tema responde a una doble motivación. Por una parte, resulta especialmente relevante por la constante evolución del marco regulatorio de la inteligencia artificial y por el esfuerzo continuado de las instituciones europeas por encontrar una respuesta jurídica adecuada a los desafíos que esta tecnología plantea. La entrada en aplicación del Reglamento (UE) 2024/1689 (AI Act), la aprobación de la nueva Directiva sobre responsabilidad por productos defectuosos y la retirada por la Comisión Europea, en 2025, de la Propuesta de Directiva sobre responsabilidad en materia de IA evidencian que la cuestión sigue abierta y que todavía se busca un modelo normativo capaz de ofrecer soluciones satisfactorias.

Por otra parte, el tema responde a una preocupación de fondo relacionada con la protección efectiva de los Derechos de las personas afectadas por sistemas de inteligencia artificial. Cuando un sistema opaco provoca un daño, la víctima puede encontrarse en una situación de especial vulnerabilidad al carecer de los medios necesarios para comprender cómo se produjo el resultado lesivo o para acreditar los hechos en un procedimiento judicial. En consecuencia, el debate sobre la responsabilidad por daños causados por la IA trasciende el ámbito puramente académico y se conecta directamente con la garantía de Derechos fundamentales y con la efectividad de la tutela judicial en una sociedad cada vez más automatizada.

El objetivo general de este trabajo es determinar si el sistema español de responsabilidad civil extracontractual y el marco regulatorio europeo que lo complementa ofrecen una respuesta suficiente y coherente frente a los daños causados por sistemas de inteligencia artificial o si, por el contrario, presentan limitaciones estructurales que exigen una adaptación normativa de mayor calado.

De este objetivo general se derivan tres objetivos específicos: primero, identificar en qué medida las características técnicas de los sistemas de IA —su autonomía, opacidad y capacidad de aprendizaje— tensionan los presupuestos clásicos de la responsabilidad civil; segundo, evaluar la adecuación de los distintos modelos de responsabilidad aplicables, desde la responsabilidad por culpa hasta la objetiva, pasando por el régimen de productos defectuosos; y tercero, examinar el régimen probatorio resultante del marco normativo europeo vigente y valorar si garantiza una tutela efectiva a las víctimas.

A estos tres objetivos se añade un cuarto, de carácter propositivo: formular propuestas concretas de reforma que, partiendo del diagnóstico anterior, ofrezcan soluciones operativas para los principales vacíos identificados.

Para responder a estos objetivos, el trabajo se articula en torno a tres preguntas fundamentales. En primer lugar, ¿puede el artículo 1902 del Código Civil, aplicado sin adaptaciones, dar respuesta a los daños causados por sistemas de IA autónomos? En segundo lugar, ¿qué modelo o combinación de modelos de responsabilidad resulta más adecuado para este tipo de daños? Y, en tercer lugar, ¿es el régimen probatorio previsto en el marco europeo suficiente para corregir la asimetría informativa estructural entre quien desarrolla y controla el sistema y quien sufre sus consecuencias?

La estructura del trabajo responde a una lógica deliberadamente progresiva, que avanza desde los fundamentos generales hacia las cuestiones más específicas y, finalmente, hacia la propuesta de reforma. Se ha optado por este orden, y no por otro, porque solo desde la comprensión previa de los presupuestos clásicos de la responsabilidad civil resulta posible apreciar con precisión en qué punto y por qué razón entran en crisis al aplicarse a la IA. Así, tras esta introducción, el capítulo segundo aborda el concepto de IA y su relevancia jurídica; el tercero examina los fundamentos del Derecho de Daños español y su confrontación con las características técnicas de estos sistemas; el cuarto analiza la tipología de daños indemnizables y los problemas de imputación; el quinto valora los distintos modelos de responsabilidad aplicables; el sexto se detiene en el régimen probatorio, identificado como el núcleo del problema; y el séptimo estudia las tendencias regulatorias en la Unión Europea, integrado por el Reglamento (UE) 2024/1689, la nueva Directiva sobre responsabilidad por productos defectuosos y el incierto futuro de la Propuesta de Directiva sobre responsabilidad en materia de IA, retirada por la Comisión en 2025.

Este último capítulo culmina, en su apartado 7.6, con tres propuestas de lege ferenda que actúan sobre los tres momentos del ciclo del daño: antes de que se produzca, en el momento en que se produce y después de que haya ocurrido. El trabajo se cierra con un apartado de conclusiones que da respuesta sistematizada a las preguntas planteadas.

Desde el punto de vista metodológico, la investigación se ha desarrollado a partir de una revisión bibliográfica de la doctrina civilista española y comparada en materia de responsabilidad civil e inteligencia artificial, completada con el análisis sistemático de las fuentes normativas relevantes, tanto nacionales (Código Civil, TRLGDCU, Ley 15/2022) como de la Unión Europea (AI Act, Directiva (UE) 2024/2853, RGPD y la retirada Propuesta de Directiva COM(2022) 496 final). Este análisis normativo se ha enriquecido con el examen de la jurisprudencia más reciente con incidencia en la materia —señaladamente la STJUE de 4 de septiembre de 2025 (asunto Quirin Privatbank)— y con el estudio de supuestos de hecho de especial relevancia práctica, como los casos Raine v. OpenAI o el episodio de clasificación errónea del sistema Google Photos, que ilustran las tensiones teóricas examinadas. La combinación de estas fuentes ha permitido transitar, de forma ordenada, desde la fijación de los objetivos hasta la formulación de las conclusiones y propuestas finales.

2. LA INTELIGENCIA ARTIFICIAL Y SU RELEVANCIA JURÍDICA

2.1. Concepto y características de la IA

La inteligencia artificial ha revolucionado, en poco más de una década, toda nuestra realidad social, económica y jurídica. Su progresiva implantación en sectores tan diversos como la medicina, el transporte, los servicios financieros o la administración de justicia ha supuesto un impacto significativo en la distinción entre las decisiones humanas y las producidas por procesos automatizados, desdibujando sus límites tradicionales de forma irreversible¹.

¹Marchena Gómez, M., *Inteligencia artificial y responsabilidad civil*, pág. 3.

Este fenómeno se explica, en gran medida por la capacidad que tiene la inteligencia artificial para automatizar tareas, optimizar recursos y realizar procesos con mayor rapidez y precisión que los seres humanos, lo que ha llevado a ceder paulatinamente espacio en la ejecución de funciones antes exclusivamente humanas, alcanzando hoy actividades que requieren un nivel de razonamiento, aprendizaje y evaluación, que el Derecho siempre había presupuesto como propias del ser humano.

No obstante, esta evolución ha transformado notablemente nuestra realidad en un muy corto periodo de tiempo y continuará haciéndolo, lo que evidencia la necesidad de establecer un marco regulatorio que permita gestionar sus efectos y que, al mismo tiempo, plantea los interrogantes fundamentales que este trabajo pretende abordar: en un entorno crecientemente automatizado, ¿a quien le corresponde asumir las consecuencias de los daños que estos sistemas causan?

La delimitación conceptual de la IA no es una tarea fácil de abordar. La doctrina ha propuesto una gran variedad de definiciones que oscilan entre aproximaciones estrictamente técnicas y caracterizaciones funcionales orientadas a sus efectos prácticos. Sin intención de agotar este debate, resulta útil poder a efectos jurídicos distinguir gracias a Pérez Luño, entre una concepción amplia que identifica la IA como el conjunto de actividades informáticas que, si fueran realizadas por el hombre, se consideran producto de su inteligencia, y una concepción más restringida centrada en los denominados sistemas expertos, esto es, programas que incorporan de manera práctica y operativa el conocimiento de un experto en una materia determinada y reproducen sus actuaciones². Esta segunda definición es la que tiene mayor relevancia en el tráfico jurídico actual.

A efectos normativos, la definición más relevante la ofrece el Reglamento (UE) 2024/1689, conocido como AI Act: un sistema de IA es un sistema basado en una máquina, diseñado para funcionar con distintos nivel de autonomía, que puede mostrar adaptabilidad después del despliegue y que para objetivos explícitos o implícitos, infiere, a partir de la información que recibe, generando resultados tales como predicciones, contenidos, recomendaciones o decisiones que pueden influir en entornos físicos o virtuales³.

² Pérez Luño, A.E., "¿Qué significa juzgar?", *Cuadernos de Filosofía del Derecho*, núm. 32, 2009, pág. 164, citado por MARCHENA GÓMEZ, M., op. cit., pág. 8.

³ 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, art. 3.1.

Esta definición, propuesta por el legislador europeo, plantea consecuencias directas para el Derecho de Daños. Un sistema capaz de influir en entornos físicos o virtuales mediante decisiones o remisiones es, por consiguiente, capaz de causar daños. Y un sistema que opera con autonomía variable es aquel cuya conducta no puede reconducirse de manera directa y unívoca a una decisión humana concreta.

La doctrina civilista ha subrayado con acierto que lo singularmente problemático de la IA no es su carácter automatizado, pues los sistemas automatizados llevan años operando sin suponer dificultades insuperables, sino su capacidad de aprendizaje y adaptación. Los sistemas basados en técnicas de aprendizaje automático, y especialmente los de aprendizaje profundo, no ejecutan instrucciones predefinidas por su programador: generan patrones de comportamiento a partir de los datos con los que interactúan, modificando su funcionamiento de manera continua y produciendo resultados que, en muchos casos, no han sido expresamente previstos por nadie⁴. Esta imprevisibilidad estructural es la que tensiona las bases sobre las que se ha construido históricamente la responsabilidad civil.

En definitiva, la IA constituye un sistema dinámico y en ocasiones opaco, cuya capacidad para generar consecuencias jurídicamente relevantes de forma autónoma impide tratarla como una mera herramienta al servicio de un operador humano que mantiene en todo momento el control efectivo sobre sus resultados. Esta característica es la que justifica, precisamente, que la atribución de responsabilidad por los daños que estos sistemas causan no pueda resolverse mediante la simple aplicación de las categorías tradicionales.

2.2. Clasificación de los sistemas de IA según el nivel de riesgo

La clasificación de los sistemas de inteligencia artificial no es solamente una cuestión descriptiva o técnica. Ya que tiene consecuencias jurídicas directas en la medida en el que el nivel de autonomía, complejidad y opacidad del sistema condiciona de forma decisiva la posibilidad de identificar el nexo causal entre su funcionamiento y el daño producido, así como la de atribuir responsabilidad a los distintos operadores que intervienen en su ciclo de vida.

⁴ Marchena Gómez, M., *op. cit.*, pág. 11.

Desde una perspectiva técnica, la doctrina distingue, con carácter general, entre la denominada inteligencia artificial débil o estrecha —que comprende la totalidad de los sistemas actuales, diseñados para tareas específicas y delimitadas— y la inteligencia artificial general, que se caracteriza por una capacidad cognitiva equiparable a la humana y por la aptitud para desempeñar cualquier tarea intelectual⁵.

La cuestión de si cabe atribuir personalidad jurídica a estos sistemas, así como sus consecuencias en materia de imputación, se examinará en apartados posteriores. El análisis debe centrarse en los sistemas de IA débil, que son los únicos operativos en el tráfico jurídico actual y, por consiguiente, los únicos que, en la práctica, pueden originar los daños que este trabajo analiza. Que se califiquen como “débil” no implica ausencia de capacidad agresiva, sino únicamente que su funcionamiento está orientado a tareas concretas y definidas, lo que no excluye que puedan generar consecuencias jurídicas de notable gravedad.

Sin embargo, la clasificación de mayor relevancia jurídica no es la anterior, sino la establecida por el AI Act, atendiendo al nivel de riesgo que cada sistema representa para los Derechos y libertades de las personas. El Reglamento consta de cuatro categorías. En primer lugar, tenemos los sistemas de riesgo inaceptable, cuyo uso queda directamente prohibido. Entre ellos se encuentran los sistemas de puntuación social de ciudadanos por parte de los poderes públicos, los sistemas de manipulación subliminal del comportamiento humano, así como determinadas aplicaciones de identificación biométrica en tiempo real en espacios públicos⁶.

En segundo lugar, los sistemas de alto riesgo, estos pueden comercializarse y utilizarse, pero están sometidos a requisitos estrictos de transparencia, trazabilidad, supervisión humana y evaluación de conformidad previas a su puesta en el mercado. El Reglamento incluye en esta categoría, entre otros, los sistemas utilizados en infraestructuras críticas, en la selección de personal y en la evaluación de estudiantes, en la administración de justicia o en la concesión de créditos⁷.

⁵ López de Mántaras, R., "Algunas reflexiones sobre el presente y futuro de la Inteligencia Artificial", *Novática*, núm. 234, 2015, págs. 97-101.

⁶ Reglamento (UE) 2024/1689, art. 5.

⁷ Reglamento (UE) 2024/1689, art. 6 y Anexo III.

En tercer lugar, tenemos los sistemas de riesgo limitado, sujetos únicamente a obligaciones de transparencia frente al usuario, como los chats o los sistemas de generación de contenido sintético.

Finalmente, los sistemas de riesgo mínimo que no están sujetos a ninguna obligación específica comprenden la mayor parte de las aplicaciones de IA actualmente en uso como los filtros de correo no deseado o los sistemas de recomendación de contenidos en plataformas de entretenimiento.

Esta clasificación por niveles de riesgo tiene una proyección directa sobre el régimen de responsabilidad civil: a mayor nivel de riesgo del sistema, mayor debe ser la protección de la víctima y menor la carga probatoria que se le impone. Como señala Solanes Llano en el marco del VIII Congreso sobre Derechos Humanos e Inteligencia Artificial, el legislador europeo ha optado por un enfoque que equilibra innovación tecnológica y protección de Derechos fundamentales, cuyo eje central es precisamente la supervisión humana de los sistemas de alto riesgo, configurada en el artículo 14 del AI Act como un requisito normativo fundamental⁸. Que ese AI Act carezca de un instrumento de responsabilidad civil complementario — tras la retirada de la Propuesta de Directiva en 2025— supone que la clasificación por riesgo existe pero carece de consecuencias reparadoras sistemáticas.

Junto a la clasificación por nivel de riesgo, es de igual importancia para el análisis probatorio la distinción que se basa en la transparencia del proceso de toma de decisiones del sistema. Desde esta perspectiva, cabe diferenciar entre los sistemas de caja blanca o transparente, en los que el proceso de decisión es rastreable y comprensible para los operadores humanos, y los sistemas de caja negra u opaca, característicos de modelos de aprendizaje profundo. En los primeros, los pasos que el algoritmo sigue son verificables y permiten reconstruir el razonamiento que conduce a un resultado determinado. En los segundos, la interacción entre múltiples capas neuronales y la complejidad del modelo matemático subyacente impiden incluso a los propios desarrolladores explicar con precisión por qué el sistema adoptó una decisión concreta en un caso específico. Esta distinción conecta directamente con los problemas probatorios que se analizarán en el apartado siguiente.

⁸ Solanes Llano, Á. y Llano Alonso, F., "Ética de la IA y Derechos Humanos", en *Actas del VIII Congreso sobre Derechos Humanos e Inteligencia Artificial*, Fundación Mainel, Valencia, 2026, p. 19.

2.3. Riesgos jurídicos derivados de su uso: el efecto caja negra

Los riesgos jurídicos que genera el uso de sistemas de IA no son hipotéticos ni futuros: se manifiestan ya en los litigios actuales, aunque la jurisprudencia todavía no haya sistematizado respuestas específicas para todos ellos. Su análisis es imprescindible para entender por qué el marco normativo vigente resulta insuficiente y porque la adaptación del Derecho de daños no debe tardar en llegar.

El primero y más significativo es el denominado efecto caja negra, que constituye el obstáculo central que la AI plantea al funcionamiento del sistema clásico de responsabilidad civil. No se trata simplemente de que el proceso de decisión sea complejo: se trata de que es estructuralmente irreproducible. Cuando un algoritmo de deep learning deniega un crédito, clasifica erróneamente a una persona por sus rasgos físicos o recomienda un tratamiento médico incorrecto, ni el perjudicado ni sus asesores jurídicos ni los propios desarrolladores del sistema pueden reconstruir el razonamiento de ese resultado. El buscar justificación a la toma de decisión puede ser algo sin respuesta, y si no puede responderse, demostrar el nexo causal entre el sistema y el daño se convierte para la víctima en una tarea prácticamente imposible⁹.

Esta obstaculización al determinar el nexo causal, que ya es de por sí uno de los elementos más complejos del Derecho de daños, hace dificultosa la acreditación de la relación de causalidad que se le exige al demandante para poder reconstruir el proceso que derivó en el daño y así poder atribuir la conducta al responsable. Pero cuando el proceso está oculto por el propio diseño de estos sistemas, la exigencia se convierte en una carga desproporcionada que, en la práctica limita el acceso a la justicia para aquellos que no poseen recursos técnicos y económicos necesarios para afrontarla.

La imposibilidad de reconstruir este proceso decisonal coloca al perjudicado en una situación notablemente desventajosa con respecto del desarrollador del sistema, comprometiendo seriamente su Derecho a la tutela judicial efectiva si el ordenamiento no articula mecanismos que corrijan esa asimetría.

⁹ Vilches, J., "IA en 'caja negra'", *Avezalia*, 25 de febrero de 2026.

Belalcazar confirma que la literatura jurídica comparada identifica de forma unánime la opacidad algorítmica y la asimetría informativa entre proveedores y usuarios como los principales obstáculos para la reparación efectiva de las víctimas¹⁰.

Un segundo riesgo es el de la imputación fallida, que se produce cuando el daño causado por un sistema autónomo no puede atribuirse a ningún sujeto concreto con la certeza probatoria que el Derecho exige. En el ciclo de vida de un sistema de IA intervienen el fabricante, el diseñador del algoritmo, quien lo entrena con datos, el operador que lo despliega y el usuario que lo emplea.

Cuando el daño deriva del autoaprendizaje del sistema, de su capacidad de aprendizaje y adaptación que lo conduce por caminos no previstos por ninguno de esos actores, la atribución individual de responsabilidad resulta problemática para todos ellos. El resultado puede ser que nadie responda, lo que la doctrina ha denominado como el problema de la decisión emergente y que contradice uno de los principios estructurales del Derecho de daños: quien crea un riesgo y se beneficia de él debe responder de sus consecuencias dañosas¹¹.

Un tercer riesgo deriva de la imprevisibilidad del comportamiento del sistema. En el Derecho de daños la previsibilidad del resultado lesivo es un elemento relevante tanto para la apreciación de la culpa como para la imputación objetiva. Los sistemas de IA basados en el aprendizaje continuo pueden generar comportamientos emergentes no previstos en el momento del diseño o del despliegue, lo que hace difícilmente aplicable el estándar clásico de previsibilidad a los supuestos en los que el daño no deriva del uso regular del sistema sino de sus propias decisiones autónomas respecto del objetivo para el que fue creado.

Existe finalmente un cuarto riesgo que merece atención específica: la discriminación algorítmica. Los sistemas entrenados con datos que reflejan sesgos históricos pueden reproducir y amplificar discriminaciones en el acceso al crédito, empleo o sanidad de forma sistemática e invisible.

¹⁰ Belalcazar Apunte, P.A., "La responsabilidad civil por daños derivados del uso de inteligencia artificial. Una revisión sistemática", *Polo del Conocimiento*, vol. 11, núm. 1, 2026, p. 814.

¹¹ Gil Nogueras, L., "La responsabilidad civil como consecuencia de daños causados por la IA", en Magro Served, V. (Coord.), *La responsabilidad civil como consecuencia de daños causados por la Inteligencia Artificial*, obra colectiva, pp. 2-3.

Su invisibilidad es precisamente lo más preocupante desde una perspectiva jurídica ya que la mayoría de los usuarios afectados no saben que han sido objeto de una decisión algorítmica, y si lo ignoran no pueden ejercer su Derechos a impugnarlas tal y como se les reconoce en el artículo 22 del RGPD¹².

Como han señalado Chaveli Donet y Merino Rus, los sesgos algorítmicos constituyen una forma de discriminación estructural cuyo origen se encuentra en los datos de entrenamiento y cuya propagación es difícil de detectar precisamente porque opera bajo la apariencia de objetividad técnica¹³.

Esta dimensión constitucional del problema exige una lectura coordinada de las normas de responsabilidad civil con el marco europeo de no discriminación, una lectura que el ordenamiento español no ha completado todavía de forma satisfactoria.

A los riesgos anteriores cabe añadir uno que opera en un plano distinto: el uso deliberadamente malicioso de sistemas de IA generativa como instrumento para causar daño a terceros. Cuando un usuario emplea un modelo generativo para crear contenido sintético fraudulento como por ejemplo, el uso de correos electrónicos falsos, facturas falsificadas, deepfakes de personas reales, el sistema no falla, funciona exactamente como debe. El problema es que el causante del daño se esconde tras un anonimato, imposible de identificar, lo que priva a la víctima de cualquier vía de reparación efectiva. Este escenario exige respuestas regulatorias específicas que no pueden alcanzarse con las herramientas clásicas de la responsabilidad civil por productos defectuosos ni por culpa.

¹² Reglamento (UE) 2016/679 del Parlamento Europeo y del Consejo, de 27 de abril de 2016, relativo a la protección de las personas físicas en lo que respecta al tratamiento de datos personales (RGPD), art. 22.

¹³ Chaveli Donet, E. y Merino Rus, R., "Sesgos, discriminación y otros algoritmos", en *Actas del VIII Congreso sobre Derechos Humanos e Inteligencia Artificial*, Fundación Mainel, Valencia, 2026, p. 32.

3. LOS FUNDAMENTOS DE LA RESPONSABILIDAD CIVIL EN EL DERECHO ESPAÑOL

3.1. Concepto y función

El Derecho de la responsabilidad civil afronta un problema que todos los ordenamientos jurídicos deben resolver: determinar quién debe soportar las consecuencias económicas de un daño cuando este no se produce en el vacío sino como resultado de la actuación de otro.

En palabras de Díez-Picazo, se trata de resolver la distribución de los infortunios: cuando ocurre un accidente del que se derivan daños a las personas o a las cosas, el ordenamiento debe decidir si quien lo experimenta no tiene otra posibilidad que la resignación o si tiene derecho a reclamar de otro su reparación¹⁴. La obligación de reparar no surge del mero hecho de que el daño se haya producido, sino de que pueda atribuirse a la conducta reprochable de un sujeto concreto. Es lo que la doctrina denomina el principio de imputación, y es el eje sobre el que gira todo el sistema.

La responsabilidad civil cumple en el ordenamiento jurídico tres funciones que conviene distinguir con claridad, porque cada una de ellas se ve afectada de manera distinta por la irrupción de la IA. La función reparadora es la principal: compensar a quien ha sufrido el daño por la actuación de otro. La función preventiva actúa de manera indirecta: quien sabe que deberá responder económicamente por los daños que cause tiene un incentivo estructural para adoptar medidas de diligencia. La función de demarcación limita los ámbitos de libertad de actuación: toda actividad que pueda generar responsabilidad constituye por ella misma un límite a la libertad de quien la desarrolla.

Esta pluralidad de funciones resulta especialmente relevante en el contexto de la IA porque el Derecho de daños no solo está llamado a reparar los daños ya producidos, sino también a configurar los incentivos bajo los que se desarrollan y despliegan estos sistemas.

¹⁴ Díez-Picazo y Ponce de León, L., *Derecho de Daños*, Civitas, Madrid, 1999, p. 43.

Como señala Belalcazar sobre la base de la literatura comparada, la responsabilidad civil funciona como una palanca de cumplimiento que incentiva la documentación, la auditoría y el monitoreo del riesgo: sin consecuencias jurídicas concretas los marcos éticos de gobernanza algorítmica, carecen de eficacia real, y los costes del daño se trasladan a las víctimas y a la sociedad en lugar de ser internalizados por los actores con capacidad de control¹⁵. Encontrar un punto de equilibrio entre la tutela efectiva de las víctimas y el fomento responsable de la innovación es precisamente el desafío central que el legislador europeo no ha resuelto de forma satisfactoria con los instrumentos actualmente vigentes.

3.2. Responsabilidad contractual y extracontractual

El ordenamiento jurídico español, siguiendo la tradición codificadora francesa distingue entre dos grandes regímenes de responsabilidad civil. La responsabilidad contractual, regulada en los artículos 1101 y siguientes del Código Civil, surge del incumplimiento o del cumplimiento defectuosos de una obligación preexistente entre las partes: el deudor de buena fe responde únicamente de los daños previsibles en el momento de contratar mientras que el deudor doloso responde de todos los daños que sean consecuencia directa e inmediata del incumplimiento, conforme a los artículos 1106 y 1107. La responsabilidad extracontractual, regulada fundamentalmente en el artículo 1902 CC, nace con independencia de cualquier vínculo previo: basta con que alguien cause un daño a otro mediante una conducta culposa para que nazca la obligación de repararlo.

Esta distinción tiene consecuencias prácticas directas en el ámbito de los daños causados por sistemas de IA. Cuando el perjudicado ha contratado con el proveedor u operador del sistema puede disponer simultáneamente de la acción contractual y de la extracontractual, lo que plantea el clásico problema del cúmulo de responsabilidades. Cuando el daños es sufrido por un tercero ajeno a cualquier relación contractual por ejemplo la víctima de un accidente causado por un vehículo autónomo o un candidato discriminado por un sistema de selección de personal, la vía disponible es la extracontractual.

¹⁵ Belalcazar Apunte, P.A., *op. cit.*, p. 817.

Junto con los daños patrimoniales y físicos, los sistemas de IA pueden generar daños morales autónomos, como la pérdida de control sobre los datos personales, cuestión sobre la que se ha pronunciado recientemente el Tribunal de Justicia de la Unión Europea y que se analiza con detalle en el apartado 4.1, al examinar la tipología de daños indemnizables.

Ambos instrumentos europeos sobre la IA— la Directiva sobre responsabilidad por productos defectuosos y la ya retirada Propuesta de Directiva de adaptación de la responsabilidad civil a la IA— fueron concebidos para actuar de forma complementaria: las normas de seguridad contribuyen a la reducción del riesgo mientras las normas de responsabilidad civil operan ex post para resarcir a las víctimas. La retirada de la segunda ha roto esa complementariedad, dejando a la Directiva de productos defectuosos como único instrumento armonizado y obligando a una lectura conjunta de lo que queda con el AI Act y las legislaciones nacionales de transposición.

3.3. Elementos de la responsabilidad civil extracontractual

El artículo 1902 del Código Civil establece que quien por acción u omisión, cause daño a otro, interviniendo culpa o negligencia, está obligado a reparar el daño causado. De este precepto la doctrina extrae cuatro presupuestos clásicos cuya concurrencia es necesaria para que nazca la obligación de indemnizar¹⁶. Hay que señalar que junto al criterio general de culpa, en nuestro sistema existe una sólida tradición de objetivización de la responsabilidad cuando la actividad desarrollada crea un riesgo que excede el actuar corriente como sucede en la conducción de vehículos a motor o en la caza.

Esta tradición resulta relevante de cara a la regulación de la IA, respecto de la cual una extensión análoga de la responsabilidad objetiva requeriría una previsión legal específica.

¹⁶ Díez-Picazo y Ponce de León, L., *op. cit.*, p. 43.

3.3.1. La acción u omisión

El primer presupuesto es la existencia de una conducta humana, activa u omisiva, causante del daño. La acción consiste en la realización positiva de un hecho que provoca el resultado lesivo. La omisión presenta una mayor complejidad dogmática, pues el deber de actuar que su infracción presupone no puede deducirse sin más del principio general de no causar daños a otro, sino que requiere una fuente específica que imponga al sujeto la obligación de intervenir para evitar el daño¹⁷.

En el contexto de los sistemas de IA, la identificación de esta conducta humana causante del daño plantea dificultades inmediatas. Cuando el daño se produce como consecuencia del funcionamiento autónomo del sistema, resulta problemático identificar la conducta humana directamente responsable del resultado lesivo. Como señala Gil Noguerras, en la IA convergen las actuaciones de múltiples agentes¹⁸. Por ello, cuando el daño deriva del autoaprendizaje del sistema, es decir, de su capacidad de tomar caminos desconocidos para cualquiera de estas personas, la imputación de una conducta concreta a cualquiera de ellos resulta problemática para todos.

La jurisprudencia comparada analizada por Sotomayor Padilla y otros, muestra una tendencia convergente a negar personalidad jurídica a la IA y buscar imputar la responsabilidad a las personas naturales o jurídicas que la diseñaron, implementaron, controlaron o utilizaron, reinterpretando los criterios de diligencia y supervisión humana para resolver disputas donde los algoritmos operan con opacidad¹⁹.

3.3.2. El daño

El daño es el presupuesto central de la responsabilidad civil y, al mismo tiempo, su límite. Sin daño no hay responsabilidad y la indemnización no puede exceder lo efectivamente sufrido. El daño indemnizable debe reunir una serie de requisitos: ha de ser cierto y efectivo, personal y antijurídico²⁰.

¹⁷ Díez-Picazo y Ponce de León, L., *op. cit.*, p. 107.

¹⁸ Gil Noguerras, L., *op. cit.*, pp. 2-3.

¹⁹ Sotomayor Padilla, P.C., Vaca Nájera, S.O., Vasconez Nolivos, D.C., Moreira Vite, M.D. y Jácome Rocha, A.I., "Jurisprudencia sobre inteligencia artificial y responsabilidad jurídica", *Ethos Scientific Journal*, vol. 4, núm. 1, 2026, p. 150.

²⁰ Díez-Picazo y Ponce de León, L., *op. cit.*, pp. 307-310.

La doctrina distingue clásicamente entre el daño patrimonial —daño emergente y lucro cesante— y el daño moral, que abarca los perjuicios no patrimoniales derivados de la lesión de bienes o intereses de la personalidad como el honor, la intimidad o la integridad psíquica.

Esta construcción clásica, pensada para perjuicios individuales, concretos y de manifestación inmediata, se ve sometida a una considerable tensión en el ámbito de los sistemas de IA, donde la amplitud y heterogeneidad de los perjuicios supera con creces lo que el legislador tuvo en mente al redactar el artículo 1902 CC. El análisis pormenorizado de esta nueva tipología de daños — físicos, patrimoniales, morales y por discriminación algorítmica—, así como de la problemática específica de los daños colectivos y difusos, se aborda de forma sistemática en el capítulo 4 de este trabajo, al que procede remitirse.

3.3.3. La relación de causalidad

La relación de causalidad constituye el nexo que debe unir la conducta del agente al daño producido. Según la teoría de la causalidad adecuada, consolidada en la jurisprudencia del tribunal supremo, no basta con que la conducta haya sido condición sine qua non del daño: es preciso además, que según un juicio de probabilidad basado en la experiencia, esa conducta resulte idónea para producir ese tipo de resultado²¹. La determinación del nexo causal debe inspirarse en la valoración de las condiciones que la buena fe señala en cada caso como índice de responsabilidad, dentro del infinito encadenamiento de causas y efectos que puede ofrecer la realidad.

Esta formulación, aparentemente flexible, choca frontalmente con la realidad de los sistemas de IA opacos, cuyo proceso decisional no puede ser reconstruido ni siquiera por sus propios creadores.

Cuando el daño se produce a través de un proceso decisional opaco, la acreditación del nexo causal se convierte en una exigencia de resultado imposible para quien no tiene acceso a la información técnica del sistema, comprometiendo seriamente su Derecho a la tutela judicial efectiva.

²¹ Díez-Picazo y Ponce de León, L., *op. cit.*, p. 338.

Como ha señalado Belalcazar sobre la base de la doctrina comparada, incluso cuando el daño es evidente, el acceso efectivo a la reparación puede verse comprometido por barreras procesales estructurales derivadas de la opacidad técnica²². Esto justifica la introducción de mecanismos de exhibición de pruebas y de presunciones de causalidad que se analizan en el capítulo sexto.

3.3.4. *La culpa o negligencia*

La culpa o negligencia constituye el criterio subjetivo de imputación sobre el que se construye el sistema del artículo 1902 del CC. En su concepción clásica, supone la infracción de un deber de diligencia: quien no actúa con el cuidado que las circunstancias exigen y de cuya omisión se deriva un daño a otro, incurre en responsabilidad²³. El parámetro referencial es el de buen padre de familia; en ámbitos especializados, el del profesional diligente, siempre adaptado a las circunstancias concretas del caso.

La jurisprudencia española ha evolucionado significativamente hacia una objetivación progresiva de la responsabilidad en los supuestos de actividades que generan riesgos extraordinarios para terceros, lo que tiende a invertir la carga de la prueba, de modo que es el demandado quien debe acreditar que actuó con la diligencia debida para quedar exonerado²⁴. Esta tendencia, desarrollada en el ámbito de los accidentes de circulación y de la responsabilidad médica, constituye un antecedente relevante para el tratamiento jurídico de la IA.

En el ámbito de los sistemas de IA, sin embargo, el propio concepto de culpa se ve profundamente trastocado. Como señala Navarro Mendizábal, cuando una IA se equivoca, no es fácil determinar si ha sido negligente, ya que no existe un estándar de diligencia aplicable a los sistemas y algoritmos²⁵.

²² Belalcazar Apunte, P.A., *op. cit.*, p. 814.

²³ Díez-Picazo y Ponce de León, L., *op. cit.*, p. 351.

²⁴ Díez-Picazo y Ponce de León, L., *op. cit.*, p. 121.

²⁵ Navarro Mendizábal, I. A. (2024). «Inteligencia artificial y responsabilidad civil», en J. A. Moreno Martínez y P. J. Femenía López (coords.), *Inteligencia artificial y Derecho de daños: cuestiones actuales*, Dykinson, 2024, p. 861.

La dificultad no reside únicamente en la prueba de la culpa, sino en la propia definición del estándar exigible en un contexto de complejidad técnica sin precedentes: los sistemas de aprendizaje profundo generan comportamientos emergentes que ningún operador humano puede prever con certeza, lo que hace que la culpa se convierta, en determinados supuestos, en un concepto jurídico de imposible aplicación práctica. La Propuesta de Directiva COM(2022) 496 final intentó resolver esta dificultad vinculando la culpa no a un estándar abstracto de diligencia, sino al incumplimiento de obligaciones concretas del AI Act²⁶. Con su retirada en 2025, ese mecanismo ha quedado sin efecto normativo, y los tribunales nacionales deben aplicar el estándar genérico a situaciones para las que no fue diseñado.

Una posible solución a esta dificultad, que se desarrolla con mayor detalle en el apartado 7.6, consiste en objetivar el estándar de diligencia a través de un sistema de certificación funcional previa: quien despliega un sistema de alto riesgo sin acreditar que comprende sus limitaciones y ha implementado los mecanismos de supervisión exigidos no podrá invocar la exoneración por diligencia debida.

Cabe añadir, además, que la generalización de la IA está modificando el propio contenido del estándar de diligencia exigible al usuario. Como apunta Abbott, una vez que los vehículos autónomos resulten más seguros que los vehículos conducidos por humanos, podría considerarse negligente conducir uno mismo en vez de utilizar un vehículo autónomo²⁷.

Esta inversión del estándar clásico, que todavía es una hipótesis doctrinal, apunta a la profundidad de la transformación que la IA está generando en los fundamentos mismos del Derecho de Daños.

Por otro lado, Pérez García ha señalado que dotar de personalidad jurídica a la IA podría parecer una solución al problema de la imputación, pero en realidad lo agrava: si detrás de la máscara jurídica sólo hay un fantasma electrónico, el proceso judicial de reclamación de responsabilidad civil se convierte en un sistema de asignación de recursos sin que ninguna persona humana responda materialmente.

²⁶ Propuesta de Directiva del Parlamento Europeo y del Consejo relativa a la adaptación de las normas de responsabilidad civil extracontractual a la inteligencia artificial, COM(2022) 496 final, art. 4.2.

²⁷ Abbott, R.B., "The reasonable computer: Disrupting the paradigm of tort liability", *George Washington Law Review*, vol. 86, núm. 1, 2018, p. 37.

La solución correcta no es crear personas jurídicas artificiales, sino articular mecanismos de imputación funcional que identifique a los operadores humanos que crearon, explotaron o controlaron el riesgo²⁸.

3.4. Las tensiones entre el modelo clásico y la IA

El análisis anterior permite identificar con precisión las incompatibilidades estructurales entre el artículo 1902 CC y los sistemas de IA avanzados. Estas tensiones no son dificultades ocasionales que puedan resolverse mediante la interpretación flexible de los preceptos vigentes, sino limitaciones estructurales que emergen de las propias características técnicas de estos sistemas.

La primera tensión afecta al presupuesto de la conducta humana. La autonomía del algoritmo provoca un efecto diluyente entre los diferentes actores que forman parte de su ciclo de vida —el fabricante, el diseñador, el programador, el formador, el operador y el usuario—, de modo que la imputación directa a cualquiera de ellos resulta problemática. Este fenómeno, que la doctrina anglosajona denomina el problema de las muchas manos, es especialmente agudo en los sistemas de aprendizaje profundo, donde el comportamiento del sistema en el momento del daño puede no ser trazable a ninguna decisión humana concreta.

La segunda tensión afecta al presupuesto de la culpa. La responsabilidad civil subjetiva requiere que el mal funcionamiento o el uso indebido de la IA pueda explicarse y atribuirse a errores humanos, pero tal nivel de explicabilidad no siempre es posible en sistemas de deep learning, donde el comportamiento algorítmico puede resultar estructuralmente opaco. Las normas nacionales de responsabilidad civil basadas en la culpa no resultan adecuadas para tramitar las reclamaciones por daños causados por productos y servicios en los que se recurre a la IA, algo que ya reconocía expresamente la Propuesta de Directiva retirada.

²⁸ Pérez García. C. , «Personalidad jurídica e inteligencia artificial», *Universitas. Revista de Filosofía, Derecho y Política*, núm. 48, 2026, pp. 89–91 .

La tercera tensión, y la más grave desde el punto de vista de la tutela del perjudicado, afecta a la prueba del nexo causal. Si se mantiene una interpretación estricta del artículo 1902 CC, existe un riesgo real de que los daños causados por sistemas de IA autónomos queden sin reparación efectiva, generando espacios de irresponsabilidad jurídica que contradicen los principios fundamentales del sistema.

La revisión sistemática de Belalcázar confirma que esta triple tensión no es una peculiaridad del ordenamiento español, sino una constante del Derecho comparado: en todas las jurisdicciones estudiadas, la aplicación rígida de la culpa puede generar vacíos de responsabilidad y dejar sin reparación efectiva a las víctimas cuando resulta imposible identificar una conducta negligente individual²⁹.

Frente a estas tensiones, el Derecho de Daños español cuenta con instrumentos que, aplicados analógicamente y desarrollados normativamente, pueden ofrecer respuestas parciales: la responsabilidad por productos defectuosos, la objetivación jurisprudencial de la responsabilidad en actividades de riesgo y los mecanismos de inversión de la carga de la prueba. Lo que resulta evidente es que la simple aplicación del artículo 1902 CC, sin adaptaciones ni complementos, resulta insuficiente para garantizar la tutela efectiva de quienes resultan perjudicados por sistemas de IA.

4. EL DAÑO CAUSADO POR SISTEMAS DE INTELIGENCIA ARTIFICIAL

4.1. Tipología de daños indemnizables

El daño es el presupuesto central y simultáneamente el límite de la responsabilidad civil: sin daño no hay obligación de indemnizar, y la indemnización no puede exceder el perjuicio efectivamente sufrido. Para que un daño sea indemnizable en el ámbito de los sistemas de IA deben concurrir las mismas condiciones que exige el ordenamiento para cualquier otro supuesto: debe estar causalmente conectado con la conducta o la actividad del responsable, ser cierto y efectivo, y recaer sobre bienes o intereses jurídicamente protegidos³⁰.

²⁹ Belalcázar Apunte, P.A., *op. cit.*, p. 813.

³⁰ Díez-Picazo y Ponce de León, L., *op. cit.*, p. 314.

Lo que cambia no es la estructura del concepto de daño, sino la amplitud y heterogeneidad de los perjuicios que estos sistemas pueden generar, que supera con creces lo que el legislador tuvo en mente al redactar el artículo 1902 CC. Esta expansión no es solo cuantitativa: los sistemas de IA generan tipologías de daño que no existían antes y que el ordenamiento clásico no está equipado para tratar sin adaptaciones.

La primera categoría comprende los daños físicos o corporales, que se producen cuando el funcionamiento defectuoso o autónomo de un sistema causa lesiones o la muerte de una persona. Los ejemplos más representativos son los accidentes causados por vehículos autónomos y los errores de diagnóstico o de tratamiento en sistemas de IA médica.

En ambos casos la víctima se enfrenta a una dificultad añadida que no existe en los accidentes tradicionales: el daño ni deriva de una conducta humana directamente identificable en el momento del resultado lesivo, sino del funcionamiento autónomo de un sistema cuya cadena de potenciales responsables puede ser extraordinariamente extensa y cuyo proceso de decisión no puede reconstruirse.

La segunda categoría abarca los daños patrimoniales, tanto el daño emergente como el lucro cesante. En este ámbito son especialmente relevantes los perjuicios derivados de decisiones algorítmicas incorrectas en la concesión de créditos, la selección de personal, la valoración de riesgos aseguradores o la gestión de inversiones financieras. La nueva Directiva (UE) 2024/2853 ha extendido expresamente el concepto de daños indemnizables para incluir también la destrucción o corrupción de datos en el ámbito doméstico, categoría especialmente relevante en contexto de la IA generativa³¹.

La tercera categoría comprende los daños morales, es decir los perjuicios no patrimoniales derivados de la vulneración de los Derechos de la personalidad. En el ámbito de la IA presentan una dimensión específica vinculada tanto a los deep fakes, que pueden vulnerar los Derechos al honor, a la intimidad y a la propia imagen reconocidos en el artículo 18 de la CE y en la Ley Orgánica 1/1982, como a los perjuicios psicológicos derivados de interacciones prolongadas con sistemas de IA generativa, como evidencian los primeros litigios internacionales en esta materia.

³¹ Directiva (UE) 2024/2853, art. 6.

A este respecto, la STJUE de 4 de septiembre de 2025 en el asunto Quirin Privatbank ha precisado que la mera pérdida de control de datos personales puede constituir un daño moral autónomo e indemnizable sin necesidad de acreditar consecuencias más graves. Como señala Oyarzún Vargas, esta jurisprudencia amplía considerablemente el ámbito de los daños morales resarcibles en el entorno digital, aunque su cuantificación práctica continúa planteando dudas en el ordenamiento español por la ausencia de criterios específicos³².

La cuarta categoría, y la más novedosa desde la perspectiva del Derecho de daños, es la de los daños por discriminación algorítmica. Se producen cuando un sistema entrenado con datos que reflejan sesgos históricos reproduce y amplifica discriminaciones en el acceso a ámbitos sensibles como el crédito, el empleo o la sanidad.

Lo más preocupante de esta categoría no es su gravedad sino su invisibilidad estructural: en la mayoría de los casos los usuarios no son conscientes de que están siendo objeto de su aceptación o rechazo es una decisión tomada por un algoritmo y si ignoran este hecho no pueden ejercer el Derecho que el artículo 22 RGPD a impugnarlas. Molina sanchez señala que la visualización de estos sesgos es una condición previa para su corrección³³.

En este sentido, herramientas como el sistema VioGen, empleado para la valoración policial del riesgo en casos de violencia de género, requieren mecanismos específicos de supervisión y auditoría que impidan que los sesgos presentes en los datos de entrenamiento se traduzcan en decisiones discriminatorias con consecuencias directas para las víctimas. La relevancia de esta cuestión ha sido objeto de debate público tras conocerse que tres de las cuatro mujeres asesinadas en la Comunidad de Madrid desde octubre de 2025 habían sido previamente clasificadas por el sistema como casos de riesgo bajo, lo que motivó peticiones institucionales de revisión de los protocolos de valoración y seguimiento del riesgo³⁴.

³² Oyarzún Vargas, F.«Consideraciones acerca de las acciones de cesación y los daños morales por la pérdida de control de los datos personales. Comentario a la sentencia del TJUE de 4 de septiembre de 2025, asunto C-655/23 (*Quirin Privatbank*)», *Cuadernos de Derecho Transnacional*, vol. 18, núm. 1, 2026, pp709-711.

³³ Molina Sánchez, M., "Mecanismos de control para prevenir los sesgos de la herramienta Viogen", en *Actas del VIII Congreso sobre Derechos Humanos e Inteligencia Artificial*, Fundación Mainel, Valencia, 2026, p. 316.

³⁴ Comunidad de Madrid. (2026, 17 de marzo). *La Comunidad de Madrid reclama la revisión de los protocolos de valoración de riesgo en el sistema VioGen2 para mujeres víctimas de violencia*.

Esta dimensión colectiva del daño algorítmico es la que mayor dificultad plantea desde el punto de vista de su reparación jurídica, pues los instrumentos clásicos de reparación del daño fueron concebidos para compensar perjuicios individuales concretos, no para remediar lesiones difusas que afectan simultáneamente a una pluralidad indeterminada de personas.

4.2. Particularidades de los daños causados por sistemas autónomos

Más allá de la tipología de daños expuesta, los sistemas de IA autónomos presentan rasgos que modifican cualitativamente la naturaleza del perjuicio causado y que no tienen equivalente en ninguna otra categoría de daños contemplada por el ordenamiento jurídico, ya que son consecuencia directa de las características técnicas que hacen a la IA jurídicamente problemática, y su comprensión es necesaria para entender por qué las respuestas normativas vigentes resultan insuficientes.

El primero de ellos es la continuidad y acumulación del daño a lo largo del tiempo. A diferencia del daño producido por un producto defectuoso tradicional, que se manifiesta en un momento concreto y puede cuantificarse, los sistemas de IA que funcionan de forma continua pueden causar perjuicios que se acumulan progresivamente sin que la víctima sea consciente de ello.

Es por ejemplo el caso de un sistema de puntuación crediticia que aplica sistemáticamente un sesgo discriminatorio durante años, impidiendo a una persona acceder a financiación en múltiples ocasiones, o en un chatbot que mantiene durante meses interacciones psicológicamente perjudiciales con un usuario vulnerable. En el caso de Adam Raine, ChatGPT detectó, mediante sus propios sistemas de moderación, 377 mensajes con contenido autolesivo del menor a lo largo de varias semanas, sin que ninguna conversación fuera interrumpida automáticamente³⁵. Esta continuidad temporal del daño plantea interrogantes específicos sobre el cómputo de los plazos de prescripción y el momento a partir del cual debe entenderse producido el daño.

³⁵ Caso *Raine v. OpenAI, Inc.*, presentado ante los tribunales federales de California en abril de 2025.

El segundo factor que debemos tener en cuenta es la escala y el efecto multiplicador del daño. Un mismo sistema de IA defectuoso o sesgado puede causar perjuicios idénticos simultáneamente a miles o incluso millones de personas, con un coste marginal prácticamente nulo para el causante. Esta asimetría entre la facilidad de causar el daño y la dificultad de su reparación individual no tiene cabida en el Derecho clásico de daños. La magnitud de estos daños, sin embargo, no se limita a los ámbitos económicos o laborales. Los sesgos algorítmicos pueden proyectarse en ámbitos de especial sensibilidad democrática, como en la formación de la opinión política.

Estudios recientes han evidenciado que los asistentes de IA actúan como intermediarios informativos no neutrales cuando los ciudadanos les formulan consultas sobre preferencias electorales, ofreciendo recomendaciones que varían sistemáticamente según la plataforma utilizada y el idioma en que se formula la pregunta, lo que pone de manifiesto que los sesgos incorporados en estos sistemas pueden tener consecuencias que trascienden la esfera individual del perjudicado y afectan al conjunto de los procesos democráticos³⁶.

Como señala Navarro Mendizábal, los daños difusos generados por decisiones algorítmicas discriminatorias se sitúan en un terreno intermedio entre el interés individual y el colectivo, pues una misma lesión puede afectar a un amplio colectivo de personas discriminadas, aunque no todas hayan llegado a sufrir el daño de forma individualizada y reclamable³⁷.

El tercer punto es la latencia del daño, esto es, la posibilidad de que los efectos lesivos del sistema no se manifiesten hasta mucho después de que el comportamiento del causante se haya producido. A diferencia de un accidente de tráfico, cuyos efectos son inmediatos y perceptibles, los daños derivados de decisiones algorítmicas en ámbitos como la salud pueden tardar años en manifestarse con claridad suficiente como para imputarlos a una causa concreta. En este sentido, la doctrina advierte que aún se desconocen los riesgos para la salud mental de los usuarios que pueden generar las aplicaciones de IA que implican la colaboración con robots o sistemas humanoides en el hogar o en entornos laborales, siendo concebible que, hoy desconocidas, se consideren daños en el futuro³⁸.

³⁶ Martínez, C., Merino Arribas, M.A. y Martínez Vallvey, F., "La Inteligencia Artificial y su impacto en el ámbito socio jurídico", *Prisma Social*, núm. 53, 2026, p. 183.

³⁷ Navarro Mendizábal, I.A., *op. cit.*, pp. 871-872.

³⁸ Quesada Páez, A., "La atribución de responsabilidad civil por actos de sistemas autónomos de inteligencia artificial: desafíos para el Derecho privado", *InDret*, núm. 1, 2026, p. 200.

Esta latencia plantea dificultades específicas respecto de los plazos de prescripción y de la determinación en la que nace el Derecho a reclamar.

El cuarto rasgo, íntimamente ligado al anterior, es la imprevisibilidad *ex ante* del comportamiento dañoso. Los sistemas de IA conectados con otros dispositivos y sistemas configuran un ecosistema cuya complejidad puede hacer que el comportamiento del sistema en el momento del daño no sea rastreable a ninguna decisión humana adoptada en el momento del despliegue.

Esta interconexión plantea interrogantes específicos sobre la aplicabilidad del estándar de diligencia exigible a los operadores y la viabilidad de las causas de exoneración basadas en el estado de la ciencia en el momento de la comercialización: si el comportamiento dañoso era estructuralmente imprevisible, el fabricante podría exonerarse invocando precisamente la misma característica técnica, la autonomía del sistema, que justificaría su mayor nivel de responsabilidad³⁹.

Junto a estos cuatro rasgos, debe mencionarse una categoría más que aún queda sin respuesta: los daños derivados del uso deliberadamente malicioso de sistemas de IA como instrumento para perjudicar a terceros.

En estos supuestos, el sistema opera conforme a su diseño y finalidad técnica; no existe un fallo de funcionamiento. Así ocurre, por ejemplo, cuando un usuario utiliza un modelo generativo para elaborar correos fraudulentos, falsificar facturas o crear *deepfakes* de personas reales. La dificultad radica en que el autor material del daño puede permanecer anónimo o resultar imposible de localizar, privando a la víctima de una reparación efectiva. Por ello, este tipo de riesgos requiere mecanismos regulatorios específicos, distintos de los instrumentos tradicionales de la responsabilidad civil, cuestión que se examina en el apartado 7.6.

³⁹ Quesada Páez, A., *op. cit.*, p. 200

4.3. Problemas de imputación objetiva y subjetiva

La imputación del daño causado por sistemas de IA presenta dificultades que operan en dos planos distintos pero interrelacionados: la determinación del sujeto al que puede atribuírsele la responsabilidad y la acreditación del nexo causal entre la conducta de ese sujeto y el resultado dañoso

4.3.1. La quiebra del estándar clásico de diligencia

El artículo 1902 CC tipifica la imputación por infracción de un deber de diligencia. Conforme a una consolidada tradición jurídica, la diligencia exigible es la ordinaria de una persona razonable, siendo la previsibilidad una condición de la responsabilidad y la imprevisibilidad un factor de exoneración. Cuando el comportamiento causante del daño se ejerce profesionalmente, la diligencia exigible es la del profesional competente en el sector correspondiente, con arreglo a las normas de cuidado que el ordenamiento establece para cada actividad⁴⁰.

Este esquema entra en crisis cuando se aplica a los sistemas de IA por una razón estructural: actualmente no existe un estándar de diligencia profesional consolidado para el desarrollo, el despliegue y la supervisión de sistemas de inteligencia artificial. La Propuesta de Directiva COM(2022) 496 final reconocía esa dificultad al vincular la culpa no a un estándar abstracto de diligencia, sino al incumplimiento de obligaciones concretas establecidas en el AI Act, tales como el desarrollo a partir de conjuntos de datos de entrenamiento que cumplan criterios de calidad, el diseño conforme a los requisitos de transparencia, la habilitación de una vigilancia efectiva por personas físicas durante el periodo de utilización o la adopción inmediata de medidas correctoras⁴¹.

En otras palabras, el legislador europeo había optado por sustituir el estándar abierto de la diligencia debida por un catálogo cerrado de obligaciones cuyo incumplimiento operaría como prueba de la culpa, lo que habría supuesto una objetivación encubierta del criterio de imputación. Con la retirada de esa Propuesta en 2025, ese mecanismo ha quedado sin efecto normativo, y la dificultad que pretendía resolver permanece intacta.

⁴⁰ Díez-Picazo y Ponce de León, L., *op. cit.*, p. 361.

⁴¹ Propuesta de Directiva IA, COM(2022) 496 final, art. 4.2.

4.3.2. Los dos escenarios de imputación causal y la insuficiencia de la teoría de la causalidad adecuada

Desde la perspectiva del nexo causal, la doctrina distingue dos escenarios con consecuencias muy distintas para la imputación. El primero es el de los daños derivados del uso regular del sistema, es decir, aquellos que se producen en el cumplimiento del objetivo primario del algoritmo. En estos casos no hay impedimento para establecer la atribución causal conforme a la teoría de la causalidad adecuada, pues se trata de un riesgo inherente a la plataforma cuya previsibilidad objetiva puede ser constatada. Pertenecen a esta categoría los daños por perfilamiento económico discriminatorio en la concesión de créditos, las vulneraciones del Derecho de propiedad intelectual derivadas del entrenamiento de modelos generativos, o los sesgos en sistemas de selección de personal: el daño específico puede no ser directamente querido por el modelo, pero la previsibilidad de que ocurra es objetivamente constatable⁴².

El segundo escenario es el de los daños producidos por el funcionamiento errático del sistema, donde los perjuicios escapan a toda posibilidad de previsibilidad porque el algoritmo, en su proceso evolutivo, se ha desviado del objetivo para el que fue creado. En este segundo escenario la causalidad adecuada resulta insuficiente: el daño escapa al pronóstico causal y no puede reconstruirse el nexo entre la conducta del responsable y el resultado lesivo conforme a los criterios de regularidad y previsibilidad objetiva⁴³.

Así puede verse en el caso de Google Photos de 2015, en el que el algoritmo de reconocimiento facial clasificó erróneamente a personas por sus rasgos étnicos sin seguir patrón lógico alguno, generando resultados que ninguno de sus diseñadores hubiera podido prever razonablemente.

⁴² Ramos Martínez, M.F., "La imputación causal en los daños derivados de la inteligencia artificial", *International Journal of Digital Law*, v. 6, e604, 2025, p. 9.

⁴³ Quesada Páez, A., *op. cit.*, p. 201.

4.3.3 Criterios causales complementarios.

Ante la insuficiencia de la causalidad adecuada en el segundo escenario, la doctrina propone aplicar criterios causales complementarios que permiten mantener la función reparadora del Derecho de Daños sin exigir una certeza causal que los sistemas opacos hacen imposible.

El primero es el modelo estadístico de la probabilidad razonable, conocido en el Common Law como “more probable than not”, que permite establecer la conexión causal no sobre la certeza de que el daño fue consecuencia de la conducta del demandado, sino sobre la probabilidad de que así fuera conforme a los datos disponibles.

El segundo criterio es el de la imputación objetiva basada en la distinción entre el riesgo permitido y el riesgo prohibido: aunque el daño concreto sea imprevisible, la autonomía del algoritmo y los riesgos que ello conlleva son conocibles por el productor, proveedor y responsable del despliegue, lo que permite imputar objetivamente el resultado lesivo en la medida en que han superado el umbral del riesgo socialmente aceptable⁴⁴.

Estos criterios complementarios guardan coherencia con el enfoque que adoptaba la Propuesta de Directiva, en la que se articulaba precisamente una presunción de causalidad basada en la probabilidad razonable de que la culpa hubiera influido en los resultados producidos por el sistema de IA, sin exigir prueba de certeza del nexo causal. No se trata de permitir al juez utilizar un estándar de prueba reducido, sino de facilitarle presumir el nexo causal entre la acción culposa y la información o la falta de información del sistema, atendiendo a las circunstancias del caso concreto⁴⁵. La retirada de esa Propuesta ha privado al sistema de ese mecanismo aunque los criterios doctrinales que la inspiraban se pueden seguir utilizando como herramienta de interpretación.

⁴⁴ Ramos Martínez, M.F., *op. cit.*, pp. 10-11.

⁴⁵ Ortiz Fernández, M., "Los sistemas de inteligencia artificial, ¿productos defectuosos?", en Moreno Martínez, J.A. y Femenía López, P.J. (Coords.), *Inteligencia Artificial y Derecho de Daños: Cuestiones Actuales*, Dykinson, Madrid, 2024, p. 918.

4.3.4. El sistema de imputación diferenciada según el riesgo

La solución más coherente no es la elección entre un modelo u otro de responsabilidad, sino la articulación de un sistema diferenciado en función del nivel de riesgo del sistema de IA. Para los sistemas de alto riesgo —los enumerados en el Anexo III del AI Act— corresponde un régimen de responsabilidad objetiva del operador, que dispense a la víctima de probar la culpa y se centre en la acreditación del daño y del nexo causal. Para los sistemas de menor riesgo, la responsabilidad puede seguir basándose en la culpa, con la condición de que el estándar de diligencia se construya a partir de las obligaciones concretas del AI Act y no de un criterio abstracto de buen padre de familia⁴⁶.

Este sistema diferenciado presenta la ventaja de equilibrar los intereses en juego: garantiza una protección reforzada a las víctimas de los sistemas más peligrosos —precisamente aquellos en los que la prueba de la culpa y del nexo causal resulta más difícil— mientras mantiene el incentivo a la diligencia en los sistemas de menor riesgo. Su principal debilidad es que la compatibilidad entre los distintos instrumentos normativos europeos que deberían articularlo no está garantizada, y que la ausencia de la Directiva específica de responsabilidad por IA hace que el sistema diferenciado exista en el plano doctrinal pero carezca todavía de plasmación normativa coherente⁴⁷.

Esta lógica de imputación diferenciada requiere, para ser operativa, que cada actor en la cadena de valor pueda ser identificado y que su participación en el ciclo de vida del sistema quede documentada. Ello conecta directamente con la propuesta de trazabilidad jurídica obligatoria que se desarrolla en el apartado 7.6.

⁴⁶ Díez Royo, M., "Cuestiones de responsabilidad civil de los sistemas de Inteligencia Artificial en las Propuestas de Directivas europeas de 28 de septiembre de 2022", *Revista de Estudios Jurídicos y Criminológicos*, núm. 8, Universidad de Cádiz, 2023, p. 261

⁴⁷ Díez Royo, M., *op. cit.*, p. 261.

5. MODELOS DE RESPONSABILIDADES APLICABLES

5.1. Las opciones regulatorias y la evolución del debate europeo

El debate sobre cómo articular la responsabilidad civil por daños causados por sistemas de IA ha girado en torno a tres grandes opciones. La primera consiste en aplicar las normas vigentes sin modificación. La segunda propone adaptar el marco existente a las características específicas de la IA. La tercera plantea la creación de un régimen jurídico nuevo y específico. La primera opción está prácticamente descartada: las normas vigentes no ofrecen respuestas adecuadas a los problemas de opacidad, autonomía y aprendizaje continuo⁴⁸. El debate se concentra entre la segunda y la tercera, con una tendencia mayoritaria hacia la adaptación del marco existente.

La revisión sistemática de Belalcázar, que analiza doce estudios de Derecho comparado, concluye que existe una convergencia doctrinal internacional hacia regímenes de responsabilidad objetiva o híbrida para los sistemas de alto riesgo, acompañados de ajustes probatorios y mecanismos de aseguramiento. Esta tendencia no responde a una ruptura con los principios del Derecho civil, sino a una adaptación coherente a nuevas fuentes de riesgo tecnológico: la responsabilidad objetiva permite redistribuir el riesgo hacia quienes se benefician económicamente del uso de la IA y se encuentran en mejor posición para prevenir el daño⁴⁹.

La evolución del legislador europeo ilustra las dificultades del debate. La Resolución del Parlamento Europeo de 2020 propuso responsabilidad objetiva para sistemas de alto riesgo y responsabilidad por culpa con inversión de la carga de la prueba para el resto. La Comisión rechazó ese enfoque en 2022 y optó por una vía más conservadora: una Directiva de mínimos con medidas principalmente procesales. Esa Propuesta fue retirada en 2025, calificada de prematura. Como han señalado Echarrí Nicolás y Fierro Rodríguez, esa narrativa no resiste un análisis riguroso: la Propuesta respondía a un vacío estructural real, y su retirada no elimina el problema sino que lo traslada a los ordenamientos nacionales sin criterios armonizados⁵⁰.

⁴⁸ Díez Royo, M., *op. cit.*, p. 254.

⁴⁹ Belalcázar Apunte, P.A., *op. cit.*, pp. 813-814.

⁵⁰ Echarrí Nicolás, L. y Fierro Rodríguez, D., "Crónica de la muerte de la Propuesta de Directiva de daños por inteligencia artificial", *El Español Jurídico (E & J)*, 5 de octubre de 2025.

5.2. La responsabilidad por culpa y sus limitaciones

El artículo 1902 CC funciona bien en los supuestos en que el daño deriva del uso regular del sistema: cuando la previsibilidad objetiva del tipo de resultado puede ser constatada. Las dificultades estructurales surgen en los supuestos de funcionamiento errático y autónomo, donde la opacidad hace imposible la identificación de una conducta humana negligente concreta.

La Propuesta de Directiva intentó construir un sustituto funcional del estándar de culpa: vincular la imputación al incumplimiento de obligaciones normativas concretas del AI Act. Esta técnica supone en la práctica una objetivización encubierta del criterio de imputación, con la ventaja de proporcionar criterios precisos y verificables. Como precisa Belalcazar, sin deberes claros de documentación, trazabilidad y control, las presunciones probatorias pierden fuerza: la responsabilidad civil debe integrarse en un ecosistema regulatorio más amplio para responder adecuadamente a los daños algorítmicos⁵¹. Su desaparición del ordenamiento europeo deja al juez nacional sin más herramienta que el estándar genérico de diligencia.

5.3. La responsabilidad por productos defectuosos

La responsabilidad por productos defectuosos constituye el modelo de referencia más próximo al que los sistemas de IA requieren, porque prescinde de la culpa del fabricante como criterio de imputación y se centra en la defectuosidad del producto. La nueva Directiva (UE) 2024/2853 introduce varias novedades decisivas. La primera es la ampliación del concepto de producto para incluir expresamente los programas informáticos y, por extensión, los sistemas de IA⁵². La segunda es la expansión del concepto de componente, que ahora incluye los servicios conexos necesarios para que el producto funcione⁵³. La tercera establece que el fabricante que diseñe un producto con capacidad de aprendizaje autónomo debe responder por los comportamientos inesperados que ese aprendizaje genere⁵⁴. La cuarta amplía el círculo de responsables potenciales a una macro categoría de operadores económicos con responsabilidad solidaria entre ellos⁵⁵.

⁵¹ Belalcazar Apunte, P.A., *op. cit.*, p. 817.

⁵² Ortiz Fernández, M., "Los daños producidos por sistemas inteligentes «defectuosos» en la Directiva europea", *Revista de Derecho Civil*, vol. XIII, núm. 1, 2026, p. 214.

⁵³ Ortiz Fernández, M., *op. cit.*, p. 215.

⁵⁴ Ortiz Fernández, M., *op. cit.*, p. 219

⁵⁵ Serra Rodríguez, A., "El ámbito subjetivo de aplicación de la Directiva (UE) 2024/2853", *Actualidad Jurídica Iberoamericana*, núm. 24, 2026, p. 49.

Serra Rodríguez ha señalado que la nueva Directiva supone una actualización necesaria del régimen de productos defectuosos, aunque plantea dudas interpretativas relevantes sobre su ámbito subjetivo de aplicación que deberán resolverse en el proceso de transposición. A diferencia de la Directiva de 1985, que protegía en la práctica al consumidor, la nueva norma extiende la protección a toda persona física independientemente de su condición de consumidor, lo que tiene importantes implicaciones para el Derecho español⁵⁶.

Sin embargo, el modelo presenta una limitación estructural: la exoneración por riesgo de desarrollo. Como ha señalado Belalcazar sobre la base del análisis de Borges, cuando el fabricante mantiene influencia significativa sobre el funcionamiento del sistema mediante actualizaciones o parámetros de diseño, resulta razonable atribuirle responsabilidad por los daños causados incluso después de la puesta en circulación; pero la causa de exoneración puede convertirse en una vía de escape cuando el estado de los conocimientos técnicos en el momento de la comercialización no permitía prever el defecto⁵⁷.

5.4. La responsabilidad objetiva en actividades de riesgo

La tradición española de responsabilidad objetiva para actividades que generan un riesgo especial —circulación de vehículos, caza, actividades nucleares— ofrece un punto de apoyo jurídico relevante para los sistemas de IA de alto riesgo. La lógica subyacente es clara: quien introduce en el tráfico jurídico una fuente de riesgo extraordinario y se beneficia de su explotación debe asumir las consecuencias dañosas, con independencia de su diligencia individual. La traslación de esta lógica a los sistemas de IA de alto riesgo es coherente con su naturaleza intrínsecamente peligrosa, y resulta especialmente adecuada cuando los principios de transparencia y aplicabilidad del sistema no pueden cumplirse⁵⁸.

Sin embargo, su extensión generalizada a todos los sistemas de IA plantea tensiones con el fomento de la innovación tecnológica. Como señala Belalcazar sobre la base de los estudios comparados de Li, Faure y Havu, la responsabilidad objetiva genera mejores incentivos preventivos desde el análisis económico del Derecho, siempre que se diseñe de manera proporcional y se complemente con mecanismos de información y supervisión⁵⁹.

⁵⁶Serra Rodríguez, A., *op. cit.*, p. 44.

⁵⁷ Belalcazar Apunte, P.A., *op. cit.*, p. 815.

⁵⁸ Quesada Páez, A., *op. cit.*, p. 203.

⁵⁹ Belalcazar Apunte, P.A., *op. cit.*, p. 815.

Por eso la solución más equilibrada es reservar la responsabilidad objetiva para los sistemas de alto riesgo del Anexo III del AI Act, manteniendo el modelo culpabilístico —con los ajustes probatorios necesarios— para el resto.

5.5. El seguro de responsabilidad civil como mecanismo complementario

Cualquiera que sea el modelo de responsabilidad que finalmente se consolide, resulta insuficiente por sí solo para garantizar la reparación efectiva de las víctimas si no va acompañado de un mecanismo de aseguramiento obligatorio. La revisión comparada de Belalcázar confirma que el seguro no sustituye a la responsabilidad civil, sino que la refuerza cuando existe un marco jurídico coherente para la gestión del riesgo algorítmico: sin reglas definidas sobre quién responde y en qué supuestos, los mercados aseguradores enfrentan dificultades para modelar el riesgo⁶⁰. Con la retirada de la Propuesta de Directiva, la evaluación diferida de la idoneidad de un seguro obligatorio que aquella preveía ha desaparecido también del horizonte regulatorio europeo.

5.6. Propuestas doctrinales de nuevos modelos

Ante la insuficiencia de los modelos existentes para responder a la integración de los daños causados por sistemas de IA, la doctrina ha formulado diversas propuestas que van más allá de la mera adaptación del Derecho vigente.

La primera propuesta doctrinal es la responsabilidad compartida en función del rol y el grado de control de cada actor en el ciclo de vida del sistema. Quesada Páez ha formulado esta propuesta con mayor precisión bajo la denominación de atribución funcional del riesgo algorítmico: lo determinante no es la conducta concreta de cada actor, sino su posición funcional en la cadena de creación, organización y explotación del riesgo; quien está en mejor posición para prevenir el daño o para asegurar sus consecuencias debe soportar la carga de la responsabilidad⁶¹.

⁶⁰ Belalcázar Apunte, P.A., *op. cit.*, p. 816.

⁶¹ Quesada Páez, A., *op. cit.*, p. 201.

La segunda propuesta es la creación de un fondo de compensación para víctimas de IA, financiado por las empresas que desarrollan o despliegan sistemas de IA, que garantice la reparación del perjudicado con independencia de las dificultades probatorias. Esta propuesta se alinea con la lógica de los fondos de garantía existentes en otros sectores de riesgo.

La tercera propuesta es la analogía con la responsabilidad por animales del artículo 1905 CC, explorada especialmente en el ámbito de los vehículos autónomos. Esta analogía ha sido criticada por su artificiosidad, pero resulta útil como argumento para justificar un régimen de responsabilidad objetiva. Como señala Pérez García, el problema no es si es posible construir ficciones jurídicas para imputar responsabilidad a sistemas artificiales: el problema es si esas ficciones son útiles y si no desvirtúan la función del Derecho al separar la responsabilidad civil de su fundamento en la persona humana⁶².

6. EL RÉGIMEN PROBATORIO: LA CARGA DE LA PRUEBA Y SUS PARTICULARIDADES

6.1. El efecto caja negra como obstáculo probatorio

El Derecho de Daños clásico descansa sobre un principio fundamental en materia probatoria: corresponde al demandante acreditar los presupuestos de la responsabilidad civil. Sin embargo, como ya se advirtió al analizar los problemas de imputación en el apartado 4.3, esta regla entra en crisis cuando se aplica a los daños causados por sistemas de IA autónomos. La opacidad técnica del proceso decisional hace que la prueba de alguno de esos presupuestos resulte objetivamente imposible o de coste desproporcionado a la víctima.

El obstáculo central que impide al demandante conocer en profundidad el proceso por el cual el daño fue causado es el efecto caja negra. En los modelos de deep learning de gran escala, el proceso de generación de resultados se desarrolla en un espacio que no es accesible ni comprensible para los propios desarrolladores⁶³. No existe un razonamiento lógico lineal al que pueda acudir para explicar ex post por que el sistema decidió de una u otra manera.

⁶² Pérez García, C., *op. cit.*, p. 89.

⁶³ Quesada Páez, A., *op. cit.*, p. 200.

La consecuencia jurídica es directa: entre la entrada de datos y la salida de la decisión se interponen capas de inferencia, correlaciones no intuitivas y parámetros no interpretables que hacen que la acreditación del nexo causal se convierta, en la práctica, en una exigencia de resultado imposible para quien no controla la información técnica del sistema.

El problema no es sólo técnico sino jurídico-procesal. Como señala Belalcazar sobre la base de Bertolini y Episcopo, la asimetría de información entre proveedores y usuarios convierte la prueba del nexo causal en una carga excesiva para los demandantes: incluso cuando el daño es evidente, el acceso efectivo a la reparación puede verse comprometido por barreras procesales estructurales⁶⁴. Frente a esta realidad, el Derecho está optando por una solución compensatoria: si la explicabilidad no es posible o no se entrega, la carga de la prueba no puede recaer íntegramente sobre quien sufre el daño.

La jurisprudencia comparada ha comenzado a articular respuestas a esta asimetría. Como señalan Sotomayor Padilla y otros, los tribunales han reinterpretado criterios como la diligencia, la relación causa-efecto, la previsibilidad del daño y la supervisión humana para resolver disputas donde los algoritmos operan con opacidad, cumpliendo una función reguladora provisional frente a la insuficiencia normativa⁶⁵. Sin embargo, esa función provisional no debería de sustituir a un marco legislativo coherente.

6.2. Los mecanismos de exhibición de pruebas como instrumento de equilibrio.

Frente a las limitaciones de los modelos de responsabilidad que se analizan en el apartado 5, la nueva Directiva sobre responsabilidad por productos defectuosos introduce un mecanismo de exhibición de pruebas que faculta al órgano jurisdiccional para ordenar al demandado que ponga a disposición del demandante la información técnica necesaria para fundamentar la reclamación. El mecanismo opera bajo estrictas condiciones: el demandante debe haber presentado hechos e indicios suficientes para garantizar la viabilidad de su pretensión y haber desplegado previamente esfuerzos razonables para obtener la información⁶⁶.

⁶⁴ Belalcazar Apunte, P.A., *op. cit.*, p. 814.

⁶⁵ Sotomayor Padilla, P.C. et al., *op. cit.*, p. 150.

⁶⁶ Ortiz Fernández, M., *op. cit.*, p. 225.

La consecuencia más relevante se produce cuando el demandado incumple la orden judicial: en este caso la Directiva establece una presunción *iuris tantum* de existencia de defecto en el producto.

Este mecanismo no trata de preparar el proceso obteniendo información sobre hechos desconocidos, sino de corregir la asimetría informativa estructural entre el perjudicado y el operador, garantizando que el Derecho a la tutela judicial efectiva no quede vacío de contenido⁶⁷.

La retirada de la Propuesta de Directiva sobre responsabilidad en materia de IA ha privado al sistema de un mecanismo de exhibición específico para los daños extracontractuales no cubiertos por la Directiva de productos defectuosos. Esa Propuesta había además previsto que el potencial demandante, al evaluar si la interposición de una demanda era para él lo más conveniente, pudiera previamente solicitar la exhibición de información para comprobar si verdaderamente debería de iniciarse un proceso, convirtiéndose así en una herramienta esencial para el acceso a la justicia. Sin ese mecanismo los daños causados por bots, sistemas de recomendación y herramientas de generación de contenido quedan huérfanos de cobertura procesal específica⁶⁸.

6.3. Las presunciones *iuris tantum* en la nueva Directiva sobre productos defectuosos

La Directiva (UE) 2024/2853 contempla dos tipos de presunciones que operan como mecanismos de alivio probatorio en favor del perjudicado. Las presunciones de defecto operan en tres supuestos: cuando el demandado incumple la orden judicial de exhibición; cuando el demandante acredita que el producto no cumple los requisitos de seguridad obligatorios; o cuando el daño proviene de un mal funcionamiento evidente durante el uso normal. Las presunciones de nexo causal operan cuando se ha acreditado tanto que el producto es defectuoso como que el daño corresponde a un tipo de perjuicio que ese defecto causa normalmente⁶⁹.

⁶⁷ Gil Noguerras, L., *op. cit.*, p. 3.

⁶⁸ Echarrí Nicolás, L. y Fierro Rodríguez, D., *op. cit.*

⁶⁹ Ortiz Fernández, M., *op. cit.*, pp. 226-228.

Para los supuestos de complejidad técnica o científica excesiva, la Directiva establece además una presunción general que permite al juez presumir tanto el defecto como el nexo causal cuando el demandante ha demostrado que el producto contribuye al daño y que es probable que sea defectuoso.

La doctrina ha señalado que estas presunciones presentan limitaciones que reducen considerablemente su eficacia práctica. Como apunta Serra Rodríguez, la nueva Directiva supone un avance innegable pero también una ocasión parcialmente desaprovechada para articular mecanismos probatorios más efectivos⁷⁰. Las exigencias acumulativas previstas son tan estrictas que, en muchos casos, el demandante habrá tenido que acreditar indicios tan próximos al propio hecho presumido que la presunción añade escaso valor probatorio real.

6.4. El régimen probatorio del AI Act y su interacción con la responsabilidad civil.

El AI Act impone a los proveedores y responsables del despliegue de sistemas de alto riesgo obligaciones con consecuencias probatorias indirectas de primera importancia: el registro de actividades para garantizar la trazabilidad, la documentación técnica verificable y las medidas de supervisión humana⁷¹.

El incumplimiento de estas obligaciones opera como prueba indiciaria de culpa del proveedor en la medida en que esas normas están específicamente destinadas a proteger frente a los daños que los sistemas de IA pueden causar. Como señala Quesada Páez, las obligaciones del AI Act revelan un juicio causal implícito sobre la necesidad de determinadas acciones para prevenir daños, de modo que no cumplirlas constituye una infracción administrativa que puede operar simultáneamente como indicio de culpa en el ámbito civil⁷².

El resultado final es un régimen probatorio fragmentado que combina la regla general del artículo 217 de la Ley de Enjuiciamiento Civil con los mecanismos de alivio de la Directiva de productos defectuosos, las normas de trazabilidad del AI Act y la inversión de la carga de la prueba de la ley 15/2022 para aquellos supuestos de discriminación algorítmica.

⁷⁰ Serra Rodríguez, A., *op. cit.*, p. 56.

⁷¹ Reglamento (UE) 2024/1689, arts. 13, 14 y 19.

⁷² Quesada Páez, A., *op. cit.*, p. 202.

La coherencia de todos estos instrumentos no está garantizada, y la ausencia de un marco armonizado beneficia estructuralmente a quienes tienen recursos para explotar las diferencias normativas entre jurisdicciones⁷³. Esta inestabilidad e insuficiencia del ordenamiento legal actual hace indispensable analizar en profundidad la arquitectura normativa de la Unión Europea y el futuro de estas regulaciones, cuestión que constituye el núcleo del siguiente apartado sobre las tendencias regulatorias en la UE.

7. TENDENCIAS REGULATORIAS EN LA UE

7.1. El marco regulatorio europeo sobre IA: estructura y lógica de conjunto

Como se anticipa en el cierre del capítulo anterior, la unión europea se ha convertido en la primera jurisdicción del mundo en contar con un marco normativo integral sobre inteligencia artificial. Este marco no responde a una norma única, sino a un conjunto de instrumentos complementarios que operan en momentos distintos del ciclo de vida de los sistemas de IA con objetivos diferentes pero íntimamente relacionados.

Como señala Díaz Galán, Europa está construyendo un modelo de IA basado en la búsqueda de un marco regulatorio ético coherente, fiable, seguro y con exigencia de responsabilidad. Este enfoque se distancia notablemente del modelo estadounidense, centrado en el mercado y la flexibilidad regulatoria, y del modelo chino orientado al control estatal. La interacción entre estos tres modelos definirá la gobernanza global de la IA, y el europeo está llamado a ser el más influyente en términos extraterritoriales debido a su carácter vinculante y garantista⁷⁴.

La distinción conceptual clave en esta arquitectura es la que separa la regulación *ex ante*, orientada a prevenir que los daños se produzcan y la de la regulación *ex post*, dirigida a garantizar la reparación cuando el daño ya se ha materializado.

⁷³ Echarrri Nicolás, L. y Fierro Rodríguez, D., *op. cit.*

⁷⁴ Díaz Galán, E.C., "La regulación jurídica de la inteligencia artificial (IA) en Europa: dos pasos decisivos, aunque insuficientes, del Consejo de Europa y la Unión Europea", Cuadernos de Derecho Transnacional, vol. 17, núm. 1, 2025, p. 388.

Como apunta Ortiz Fernandez, estas dimensiones son las dos caras de una misma moneda: las normas sobre seguridad contribuyen a reducir el riesgo, mientras que las normas sobre responsabilidad civil operan tras el siniestro para resarcir a las víctimas⁷⁵. Lamentablemente, con la retirada de la Propuesta de Directiva en 2025, esta necesaria articulación jurídica se ha visto parcialmente fracturada.

7.2. Reglamento europeo de Inteligencia Artificial (AI Act)

El Reglamento (UE) 2024/1689, en vigor desde el 1 de agosto de 2024, constituye el primer instrumento de regulación integral de la IA con carácter vinculante a nivel mundial. Su enfoque se basa esencialmente en el riesgo: a mayor nivel de riesgo del sistema mayor es la intensidad de las obligaciones impuestas. Para los sistemas clasificados como de alto riesgo, el AI Act impone un conjunto exhaustivo de deberes que abarcan la gestión de riesgos, la gobernanza de datos, la documentación técnica, la transparencia, la supervisión humana y la robustez del sistema⁷⁶.

Como subraya Toribio Carrera, el Reglamento establece normas de transparencia que exigen a los proveedores diseñar y desarrollar los sistemas de alto riesgo garantizando un nivel suficiente de explicabilidad, de modo que los responsables del despliegue puedan interpretar y utilizar correctamente sus resultados de salida⁷⁷. Sin embargo, la naturaleza del AI Act es fundamentalmente preventiva: su objetivo principal es reducir la probabilidad de que los daños ocurran, no establecer mecanismos de compensación patrimonial. Si bien es cierto que sus obligaciones de trazabilidad y documentación tienen consecuencias probatorias indirectas muy relevantes, por sí solas resultan insuficientes para garantizar la reparación efectiva de las víctimas en sede civil.

Cabe destacar que la aplicación de AI Act es gradual y escalonada. Aunque algunas obligaciones ya son plenamente exigibles, otras de carácter más técnico entrarán en vigor en los próximos años. Esta aplicación progresiva, diseñada para otorgar un margen de adaptación a los operadores, es susceptible de generar un periodo transitorio de inseguridad jurídica en el tráfico.

⁷⁵ Ortiz Fernández, M., *op. cit.*, p. 203.

⁷⁶ Reglamento (UE) 2024/1689, arts. 6 y Anexo III.

⁷⁷ Toribio Carrera, Á., "El marco regulatorio europeo de la inteligencia artificial y su impacto en el sistema judicial español", *Revista de Estudios Europeos*, núm. 87, 2026, p. 572.

7.3. La nueva Directiva sobre responsabilidad por productos defectuosos

La Directiva (UE) 2024/2853, aprobada el 23 de octubre de 2024, se erige como el primer instrumento europeo de responsabilidad civil que aborda frontalmente los daños causados por productos que incorporan sistemas de IA. Sus principales implicaciones dogmáticas han sido objeto de análisis detallado en el capítulo quinto de este trabajo.

Desde una óptica estrictamente regulatoria, su novedad más trascendente es la adopción de un enfoque de armonización máxima. A diferencia de la ya retirada Propuesta de Directiva sobre responsabilidad en materia de IA, que optaba por un enfoque de mínimos, esta nueva norma garantiza un nivel uniforme de protección en toda la Unión Europea.

Su transposición al ordenamiento jurídico español, que deberá completarse antes del 9 de diciembre de 2026, exigirá la modificación sistemática de los artículos 128 y siguientes del TRLGDCU o en su defecto la aprobación de una ley especial. Como advierte Serra Rodríguez, esta segunda vía resulta preferible desde el punto de vista de la técnica legislativa, dado que la Directiva amplía su paraguas protector a toda persona física, con independencia de que ostente o no la condición general de consumidor⁷⁸.

7.4. La Propuesta de Directiva sobre responsabilidad en materia de IA: estado actual y perspectivas

La Propuesta COM (2022) 496 final, fue formalmente retirada por la Comisión en 2025, respaldada por la Comisión de Mercado Interior del Parlamento Europeo, que la tildó de prematura, innecesaria y perjudicial para el desarrollo de la innovación. Esta decisión institucional exige una valoración crítica firme, que constituye uno de los ejes de este trabajo: la propuesta no representaba una duplicidad normativa innecesaria, sino la respuesta ineludible a un vacío estructural que el acervo vigente es incapaz de cubrir. Su retirada deja en un limbo jurídico la cobertura de los daños extracontractuales causados por chatbots, sistemas de recomendación, herramientas de generación de contenido y en general cualquier sistema algorítmico que opere bajo la lógica de prestación de servicios y no como un producto en sentido estricto.

⁷⁸ Serra Rodríguez, A., *op. cit.*, p. 44

Como argumentan con acierto Echarri Nicolas y Fierro Rodríguez, la narrativa oficial que ha justificado esta retirada no resiste un análisis técnico riguroso. Las reglas procesales de inversión de la carga de la prueba y el acceso a fuentes de prueba son mecanismos plenamente consolidados en otros sectores del Derecho de la unión, como el Derecho de la Competencia, el consumo o la protección de datos, donde han demostrado una notable eficacia para corregir asimetrías estructurales sin provocar una litigiosidad abusiva⁷⁹. A ello se le suma una innegable dimensión política: la intensa presión ejercida por las grandes corporaciones tecnológicas sobre el proceso regulatorio europeo, ampliamente documentada en la doctrina académica. El éxito de estos lobbies en la paralización de la norma resulta especialmente alarmante en la medida en que las víctimas de daños algoritmos carecen de la capacidad de influencia y organización que ostentan los operadores de la industria⁸⁰.

7.5. La adaptación del Derecho de Daños al contexto tecnológico: valoración crítica

El análisis integral del marco regulatorio europeo revela con claridad tanto sus fortalezas como sus carencias estructurales. En el activo de este balance debe situarse el reconocimiento institucional explícito de que las normas generales de responsabilidad civil han quedado obsoletas frente al fenómeno tecnológico. Resultan igualmente positivos la expresa ampliación del concepto de producto, para albergar al software, la expansión del círculo de sujetos responsables y la introducción de los ya citados mecanismos de alivio probatorio.

Asimismo, el AI Act aporta un indudable valor al fijar un estándar normativo de diligencia para los operadores de sistemas de alto riesgo, cuyo incumplimiento acarrea consecuencias administrativas con un innegable potencial de arrastre a la jurisdicción civil.

Por el contrario, en el pasivo debe de anotarse la fragmentación normativa, fruto de la coexistencia de instrumentos jurídicos con distintos ámbitos de aplicación, grados de armonización y regímenes de responsabilidad.

⁷⁹ Echarri Nicolás, L. y Fierro Rodríguez, D., *op. cit.*

⁸⁰ Echarri Nicolás, L. y Fierro Rodríguez, D., *op. cit.*

La retirada de la Propuesta de Directiva sobre responsabilidad en materia de IA ha dejado sin respuesta los interrogantes nucleares del debate jurídico: cómo imputar a la responsabilidad cuando el sistema opera con un grado de autonomía imprevisible, como sortear procesalmente la barrera probatoria de la caja negra y en definitiva, cómo evitar que las víctimas de decisiones algorítmicas lesivas queden un en una situación de desamparo material.

En el ámbito nacional, la necesaria adaptación del Derecho de Daños español deberá articularse en dos niveles superpuestos. En primer lugar, mediante una transposición rigurosa y meditada de la Directiva sobre productos defectuosos antes de diciembre de 2026. En segundo lugar, a través de una aplicación integradora y sistemática del acervo disponible (el AI Act, la nueva Directiva y la ley 15/2022 sobre igualdad de trato). Este esfuerzo hermenéutico permitirá, aun en ausencia de la frustrada Directiva específica, construir un sistema de tutela jurisdiccional más sólido que el que puede ofrecer el propio artículo 1902 del Código Civil operando en solitario.

Tal y como concluye Quesada Paez, el propio fundamento dogmático de la culpa se encuentra profundamente tensionado por la IA. Ante la inexistencia de un estándar de diligencia unívocamente aplicable a la arquitectura algorítmica, pretende resolver estos conflictos mediante la mera aplicación literal del artículo 1902 CC, sin recurrir a adaptaciones o complementos normativos, resulta a todas las luces insuficiente para garantizar el Derecho a una tutela judicial efectiva de los perjudicados⁸¹.

7.6. Propuestas de lege ferenda: hacia un sistema integrado de prevención, identificación y reparación

El análisis del marco regulatorio europeo desarrollado en los epígrafes precedentes permite extraer una conclusión ineludible: el sistema vigente sufre de una carencia estructural. Regula la IA desde una perspectiva preventiva, pero fracasa al no garantizar la reparación efectiva de los daños que esta tecnología produce.

⁸¹ Quesada Páez, A., *op. cit.*, p. 201.

El AI Act impone obligaciones ex ante a los operadores de sistemas de alto riesgo; la Directiva de productos defectuosos cubre un espectro relevante pero limitado; y vacío dejado por la Propuesta de Directiva sobre responsabilidad en materia de IA ha dejado en el aire las incógnitas jurídicas más complejas.

Frente a este diagnóstico, y como aportación central de este trabajo al debate doctrinal, se formulan a continuación tres propuestas de reforma que, articuladas en su conjunto, construyen un sistema integrado capaz de operar sobre los tres momentos críticos del ciclo del daño: antes de que se materialice, en el instante en que se produce y una vez que el siniestro ha tenido lugar.

Estas tres propuestas no operan como alternativas excluyentes, sino como mecanismos complementarios: la primera interviene en el plano de la prevención, la segunda en el de la identificación del causante y la tercera en el de la imputación y reparación. Es precisamente esta coherencia interna la que dota al conjunto de la solidez para erigirse como una auténtica propuesta de lege ferenda.

A) Primera propuesta: la certificación de aptitud funcional como requisito habilitante para el despliegue de sistemas de alto riesgo

Esta primera medida parte de una constatación desarrollada en el capítulo tercero: en la actualidad, no existe un estándar de diligencia unívoco y aplicable al desarrollo, despliegue y supervisión de sistemas de IA. La retirada Propuesta de Directiva intentó erigir dicho estándar vinculando la culpa al incumplimiento de obligaciones concretas del AI Act, pero su desaparición ha dejado ese parámetro sin traducción normativa en la esfera de la responsabilidad civil. La propuesta que aquí se defiende va un paso más allá: convertir la acreditación de ese estándar en un requisito habilitante, previo e inexcusable, para el despliegue de sistemas de alto riesgo.

En concreto, se propone instaurar un sistema de certificación de aptitud funcional obligatoria para todo operador que pretenda desplegar los sistemas catalogados en el Anexo III del AI Act.

Esta certificación no exigiría una comprensión técnica exhaustiva del código fuente o del funcionamiento algorítmico interno —algo materialmente imposible en los modelos de deep learning, incluso para sus creadores-, sino la acreditación fehaciente de tres elementos funcionales: que el responsable del despliegue comprende qué tipo de decisiones genera el sistema y en qué contextos es susceptible de fallar; que ha implementado de manera efectiva los mecanismos de supervisión humana exigidos por el artículo 14 del AI Act; y que dispone de un protocolo de actuación inmediato ante resultados imprevistos o lesivos.

La trascendencia de esta propuesta en el ámbito de la responsabilidad civil radica en su conexión directa con el régimen de imputación. El operador que despliegue un sistema de alto riesgo careciendo de esta certificación —o que, poseyéndola, incumpla materialmente sus condiciones— no estaría legitimado para invocar la exoneración por diligencia debida. La certificación operaría, así, como un estándar objetivado de culpa: su ausencia o contravención constituirá prueba de presunción de negligencia en el despliegue, aliviando de forma drástica la pesada carga probatoria que hoy soporta la víctima.

Esta medida enlaza con el artículo 4 del AI Act, que ya impone a los proveedores y responsables garantizar un nivel suficiente de "alfabetización en materia de IA" entre su personal. El salto cualitativo consiste en dotar a esa obligación administrativa de consecuencias patrimoniales directas, transformando un deber de cumplimiento normativo en un criterio dogmático de imputación civil.

Naturalmente, la implementación de esta propuesta entraña desafíos técnicos —como la definición de los estándares por parte del Comité Europeo de IA— y el riesgo de generar cargas burocráticas. Sin embargo, el ordenamiento jurídico ya ha integrado con éxito habilitaciones análogas en sectores de riesgo comparable (como las licencias de conducción, las habilitaciones médicas o el manejo de mercancías peligrosas). La IA de alto riesgo exige y merece idéntico tratamiento.

B) Segunda propuesta: la identificación obligatoria del usuario como condición de acceso a sistemas de IA generativa

Esta segunda vía aborda un vacío material distinto. Mientras que la certificación funcional responde a los daños originados por un fallo intrínseco del sistema, la identificación obligatoria ataca los daños causados deliberadamente por un usuario que instrumentaliza la IA para perjudicar a un tercero.

Cuando un individuo emplea un modelo generativo para pergeñar un correo electrónico fraudulento (phishing), falsificar documentación, crear un deepfake lesivo o usurpar una identidad, el sistema no está fallando: opera exactamente conforme a su diseño. El factor determinante es el dolo del usuario. En estos supuestos, derivar la responsabilidad hacia el fabricante resulta jurídicamente insostenible, pues el uso intencionadamente malicioso rompe el nexo causal. El responsable directo es el usuario; sin embargo, en la práctica, este suele operar bajo el amparo del anonimato, desde jurisdicciones remotas o mediante perfiles opacos. Sin la identificación del causante material, la víctima queda desprovista de tutela judicial y el sistema de responsabilidad civil se torna inoperante.

Por ello, se propone exigir la identificación mediante un documento oficial verificado como condición sine qua non para acceder a sistemas de IA generativa y a aquellas herramientas cuyo output (texto, imagen, audio o video sintético) posea un alto potencial lesivo. Esta identificación no tendría carácter público: quedaría custodiada por el proveedor bajo estrictas obligaciones de confidencialidad —similares a las que rigen la retención de datos de tráfico en las telecomunicaciones—, siendo accesible de forma exclusiva mediante mandamiento judicial en el seno de un procedimiento civil o penal.

La tensión innegable de esta medida con el Derecho a la privacidad y el RGPD puede resolverse apelando al artículo 6.1.c (cumplimiento de una obligación legal) o al 6.1.f (interés legítimo). La injerencia resultaría plenamente proporcional gracias a la reserva jurisdiccional para el acceso a los datos y a la modulación del requisito según el nivel de riesgo de la herramienta. Aunque la extraterritorialidad plantee problemas prácticos frente a proveedores globales, el peso del mercado europeo (reforzado por el efecto expansivo o efecto Bruselas del RGPD y el AI Act) hace que esta exigencia sea difícilmente eludible, permitiendo reducir significativamente la impunidad del anonimato.

C) Tercera propuesta: la limitación de responsabilidad diferenciada según el rol en la cadena de valor y la trazabilidad jurídica obligatoria

El tercer planteamiento aborda el que es, dogmáticamente, el mayor escollo de la responsabilidad civil algorítmica: la distribución de la carga resarcitoria entre la multiplicidad de actores que intervienen en el ciclo de vida de la IA, especialmente cuando el fallo se produce en un punto de la cadena técnica imposible de aislar.

Se propone, por tanto, instaurar un régimen de limitación de responsabilidad diferenciada que atienda al rol funcional de cada actor, anudado inexcusablemente a una obligación de "trazabilidad jurídica" de las decisiones. La responsabilidad se fragmenta entre el sujeto que genera los datos de entrenamiento, el que diseña el modelo, el que lo comercializa y el que finalmente lo despliega. Cada operador respondería en estricta proporción a su grado de control sobre el riesgo y al beneficio patrimonial extraído de su explotación, abrazando la tesis de la "atribución funcional del riesgo algorítmico" defendida por Quesada Páez⁸².

El punto central de esta distribución es la trazabilidad jurídica obligatoria. No se exige la trazabilidad técnica del razonamiento algorítmico, sino el registro auditable del ciclo de vida del sistema: qué actor tomó qué decisión, qué bases de datos se inyectaron, qué actualizaciones se ejecutaron y qué protocolos de supervisión humana se aplicaron en cada fase. En el plano preventivo, esta exigencia documental genera formidables incentivos para la diligencia profesional.

En el plano reparador, dota a la autoridad judicial de los elementos fácticos necesarios para ubicar el fallo en un eslabón concreto de la cadena, exonerando a la víctima de la prueba cuasi imposible de la culpa individual.

Debe admitirse, que el avance hacia sistemas dotados de autonomía emergente o Inteligencia Artificial General (AGI) podría acabar difuminando por completo esta noción de "cadena de valor". En ese escenario futuro, resultaría fundamental avanzar hacia modelos de responsabilidad colectiva sufragados mediante fondos de garantía sectoriales.

⁸² Quesada Páez, A., *op. cit.*, p. 201.

No obstante, para la realidad tecnológica actual —sistemas operados, entrenados y comercializados por entes concretos— esta trazabilidad es plenamente operativa y urgente.

D) Consideración de conjunto: las tres propuestas como sistema integral

El engranaje de estas tres propuestas articula un sistema coherente. La certificación de aptitud funcional actúa antes del uso, encajando la formación y supervisión en un criterio de imputación. La identificación obligatoria actúa durante el uso, neutralizando el anonimato doloso. La trazabilidad jurídica actúa después del daño, garantizando su correcta imputación en cadenas de valor complejas.

Por separado, ninguna de estas medidas posiblemente sea la correcta. En conjunto, sin embargo, cubren los principales vacíos indemnizatorios de la IA moderna con plena proporcionalidad, sin asfixiar la innovación tecnológica ni restringir usos socialmente valiosos. En definitiva, este trabajo sostiene que, si bien el reto de la responsabilidad civil algorítmica no admite soluciones a día de hoy definitivas, si exige decisiones normativas valientes en el presente. La inacción regulatoria no es una postura neutral: constituye una decisión política y jurídica que traslada de forma silenciosa el coste de los daños tecnológicos desde los operadores que se lucran con ellos hacia las víctimas que los padecen. Una externalización del riesgo que el Derecho de Daños no puede ni debe tolerar

8. CONCLUSIONES

El objetivo de este trabajo era determinar si el sistema español de responsabilidad civil extracontractual y el marco regulatorio europeo que lo complementa ofrecen una respuesta suficiente frente a los daños causados por sistemas de inteligencia artificial. El análisis realizado permite extraer las siguientes conclusiones.

Primera. El artículo 1902 CC, aplicado sin adaptaciones, resulta estructuralmente insuficiente. Y ello no porque los daños que la IA causa sean de distinta naturaleza que los ocasionados por otros actores —siguen siendo lesiones de bienes jurídicamente protegidos que merecen reparación—, sino porque las condiciones en que se producen no encajan en los presupuestos clásicos. Lo que cambia no es la naturaleza del daño, sino la posibilidad de atribuir su origen a una conducta humana concreta. Es precisamente esa dificultad la que convierte a la IA en un auténtico desafío para el Derecho de Daños.

Segunda. Las tensiones sobre el artículo 1902 CC operan en tres planos que exigen respuestas diferenciadas. El primero afecta a la conducta humana: la autonomía del sistema diluye la imputación entre tantos actores que resulta casi imposible señalar a uno solo como responsable. El segundo afecta a la culpa: al no existir un estándar de diligencia aplicable a los algoritmos, el único camino viable es sustituirlo por un catálogo cerrado de obligaciones concretas cuyo incumplimiento opere como prueba de negligencia. El tercero afecta al nexo causal: el efecto caja negra convierte la prueba de la causalidad en una exigencia imposible para la víctima, lo que hace imprescindible introducir mecanismos que alivien esa carga. Ninguno de estos tres problemas puede resolverse mediante una interpretación más flexible del precepto.

Tercera. Ningún modelo de responsabilidad funciona por sí solo; todos necesitan complementarse. La responsabilidad por culpa resuelve bien los casos en que el daño deriva del uso regular del sistema, pero queda sin respuesta cuando el algoritmo genera comportamientos imprevisibles o inexplicables. La responsabilidad por productos defectuosos ofrece un punto de partida más sólido al prescindir de la culpa, pero la exoneración por riesgo de desarrollo amenaza con convertirse, dado el ritmo de evolución de estos sistemas, en una vía de escape para los fabricantes.

Y la responsabilidad objetiva en actividades de riesgo, aun siendo la solución más coherente para los sistemas de alto riesgo, genera —si se extiende de forma generalizada— tensiones con el fomento de la innovación que el legislador europeo no ha podido aún resolver. Cada modelo aporta algo, pero ninguno cubre por sí solo todo el abanico de daños que la IA puede ocasionar.

Cuarta. El régimen probatorio constituye el problema más urgente. Quien sufre el daño no tiene acceso a la información necesaria para probarlo, mientras que quien la posee carece de incentivo para compartirla. Esta asimetría hace que la regla general de la carga de la prueba opere, en la práctica, como una barrera que impide a muchas víctimas llegar a un juicio. Los mecanismos de exhibición de pruebas y las presunciones *iuris tantum* de la nueva Directiva de productos defectuosos apuntan en la dirección correcta, pero las condiciones acumulativas que exigen para activarse limitan severamente su eficacia.

El avance necesario es la inversión de la carga de la prueba, al menos para los sistemas de alto riesgo, siguiendo la línea que ya marca la Ley 15/2022 en materia de discriminación algorítmica. Es aquí donde el Derecho puede intervenir con mayor eficacia, sin necesidad de rediseñar desde cero las categorías sustantivas de la responsabilidad civil.

Quinta. El marco europeo vigente no ha resuelto todavía el problema. El AI Act es una herramienta preventiva valiosa, pero prevenir y reparar son cosas distintas: el Reglamento reduce la probabilidad de que los daños ocurran, pero no resuelve en gran medida los problemas cuando finalmente se producen.

La nueva Directiva sobre responsabilidad por productos defectuosos cubre un espacio importante, pero deja fuera una amplia categoría de daños causados por sistemas que no encajan en el concepto técnico de producto —chatbots, sistemas de recomendación o herramientas de generación de contenido—. La retirada de la Propuesta de Directiva sobre responsabilidad en materia de IA en 2025 ha dejado ese hueco vacío; una decisión que no respondió únicamente a razones jurídicas, sino que pudo verse influida por la presión de las grandes corporaciones tecnológicas que lograron presentar la regulación como un obstáculo a la innovación.

Sexta. Como respuesta al cuarto objetivo, se propone un sistema integrado de reforma articulado en tres medidas complementarias. La certificación de aptitud funcional actúa antes del uso, integrando la formación y la supervisión en criterio de imputación; la identificación obligatoria del usuario actúa durante el uso, neutralizando el anonimato doloso en el empleo de la IA generativa; y la trazabilidad jurídica obligatoria, unida a un régimen de responsabilidad diferenciada según el rol en la cadena de valor, actúa después del siniestro, garantizando la correcta imputación del daño. Por separado, ninguna es la ideal pero en conjunto, podrían cubrir los principales vacíos indemnizatorios con plena proporcionalidad, sin asfixiar la innovación ni restringir usos socialmente valiosos.

Séptima. El reto de la responsabilidad civil algorítmica no admite soluciones definitivas, pero sí exige decisiones normativas que deben tomarse cuanto antes. Será imprescindible seguir de cerca el instrumento regulatorio que finalmente sustituya a la Propuesta retirada, así como la jurisprudencia que los tribunales españoles y europeos vayan construyendo en aplicación del marco vigente, lo que abre una línea de investigación que este trabajo no puede agotar. En el fondo, el problema analizado trasciende lo jurídico y es un problema de poder: los sistemas de IA permiten a quienes los desarrollan adoptar decisiones que afectan diariamente a miles de personas sin que estas puedan cuestionarlas.

El Derecho de Daños, con sus funciones reparadora, preventiva y de demarcación, es el instrumento más adecuado para corregir ese desequilibrio, pero solo si está dispuesto a renovarse, como ya lo hizo ante los riesgos de la Revolución Industrial. La inacción regulatoria no es neutral: traslada silenciosamente el coste de los daños tecnológicos desde quienes se lucran con ellos hacia las víctimas que los padecen. Una externalización del riesgo que el Derecho de Daños no puede ni debe tolerar, porque su coste no se mide en términos abstractos, sino en Derechos fundamentales de personas que hoy carecen de protección.

BIBLIOGRAFÍA

1. LEGISLACIÓN

Real Decreto de 24 de julio de 1889, por el que se publica el Código Civil. Gaceta de Madrid núm. 206, de 25 de julio de 1889.

Ley Orgánica 1/1982, de 5 de mayo, de protección civil del Derecho al honor, a la intimidad personal y familiar y a la propia imagen. BOE núm. 115, de 14 de mayo de 1982.

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. DOCE L 210, de 7 de agosto de 1985.

Reglamento (UE) 2016/679 del Parlamento Europeo y del Consejo, de 27 de abril de 2016, relativo a la protección de las personas físicas en lo que respecta al tratamiento de datos personales y a la libre circulación de estos datos (RGPD). DOUE L 119, de 4 de mayo de 2016.

Ley 15/2022, de 12 de julio, integral para la igualdad de trato y la no discriminación. BOE núm. 167, de 13 de julio de 2022.

Propuesta de Directiva del Parlamento Europeo y del Consejo relativa a la adaptación de las normas de responsabilidad civil extracontractual a la inteligencia artificial, COM(2022) 496 final, de 28 de septiembre de 2022.

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 (AI Act). DOUE L, de 12 de julio de 2024.

Directiva (UE) 2024/2853 del Parlamento Europeo y del Consejo, de 23 de octubre de 2024, sobre responsabilidad por los daños causados por productos defectuosos. DOUE L, de 18 de noviembre de 2024.

2. JURISPRUDENCIA

Tribunal Supremo (Sala de lo Civil), Sentencia de 30 de junio de 1959.

Tribunal de Justicia de la Unión Europea, Sentencia de 4 de septiembre de 2025, asunto C-655/23, *Quirin Privatbank*, ECLI:EU:C:2025:655.

Caso *Raine v. OpenAI, Inc.*, Tribunal Federal del Distrito Norte de California, presentado en abril de 2025.

3. OBRAS DOCTRINALES

Abbott, R. B. (2018). The reasonable computer: Disrupting the paradigm of tort liability. *George Washington Law Review*, 86(1).

Astray Chacón, M. P. (s.f.). Daños causados por la inteligencia artificial. APM 3.9. Asociación Profesional de la Magistratura.

Belalcazar Apunte, P. A. (2026). «La responsabilidad civil por daños derivados del uso de inteligencia artificial. Una revisión sistemática». *Polo del Conocimiento*, vol. 11, núm. 1, pp. 800-820.

Chaveli Donet, E. y Merino Rus, R. (2026). «Sesgos, discriminación y otros algoritmos». En: *Actas del VIII Congreso sobre Derechos Humanos e Inteligencia Artificial*. Valencia: Fundación Mainel, p. 32.

Díez-Picazo y Ponce de León, L. (1999). *Derecho de daños*. Civitas.

Díaz Galán, E. C. (2025). «La regulación jurídica de la inteligencia artificial (IA) en Europa: dos pasos decisivos, aunque insuficientes, del Consejo de Europa y la Unión Europea». *Cuadernos de Derecho Transnacional*, vol. 17, núm. 1, pp. 380-400.

Díez Royo, M. (2023). «Cuestiones de responsabilidad civil de los sistemas de Inteligencia Artificial en las Propuestas de Directivas europeas de 28 de septiembre de 2022». *Revista de Estudios Jurídicos y Criminológicos*, núm. 8. Universidad de Cádiz, pp. 245-270.

Echeverría Alemán, D., *Inteligencia Artificial y Estado de Derecho en 2026: Transformaciones Regulatorias, Fundamentos Normativos y Reconfiguración del Estado de Derecho Algorítmico*, The Amsterdam Institute, Madrid, 2026.

Evangelio Llorca, R. (2024, 7 de noviembre). Responsabilidad civil por daños causados por IA [Ponencia]. Ciclo El Impacto del Reglamento de Inteligencia Artificial: De la Transparencia Algorítmica a la Responsabilidad Civil, Universidad de Valladolid.

García-Chamón Cervera, E. (s.f.). La responsabilidad civil como consecuencia de daños causados por la IA. En V. Magro Servet (Coord.), *La responsabilidad civil como consecuencia de daños causados por la Inteligencia Artificial*.

Gil Nogueras, L. (s.f.). La responsabilidad civil como consecuencia de daños causados por la IA. En V. Magro Servet (Coord.), *La responsabilidad civil como consecuencia de daños causados por la Inteligencia Artificial*.

López de Mántaras, R. (2015). Algunas reflexiones sobre el presente y futuro de la Inteligencia Artificial. *Novática*, (234).

López Mas, P. J. (2024). La incidencia en la conducción automatizada de la nueva Propuesta de Directiva sobre responsabilidad civil en materia de inteligencia artificial, y sus evidentes disfuncionalidades. En J. A. Moreno Martínez y P. J. Femenía López (Coords.), *Inteligencia artificial y Derecho de daños: cuestiones actuales*. Dykinson.

Marchena Gómez, M. (s.f.). *Inteligencia artificial y responsabilidad civil*.

Martínez, C., Merino Arribas, M. A. y Martínez Vallvey, F. (2026). «La Inteligencia Artificial y su impacto en el ámbito socio jurídico». *Prisma Social*, núm. 53.

Molina Martínez, L. (2024). Transparencia algorítmica y responsabilidad civil en el marco de la intermediación de las plataformas de alojamiento turístico. En J. A. Moreno Martínez y P. J. Femenía López (Coords.), *Inteligencia artificial y Derecho de daños: cuestiones actuales*. Dykinson.

Molina Sánchez, M. (2026). «Mecanismos de control para prevenir los sesgos de la herramienta Viogen». En: *Actas del VIII Congreso sobre Derechos Humanos e Inteligencia Artificial*. Valencia: Fundación Mainel, p. 316.

Moreno Martínez, J. A. y Femenía López, P. J. (Coords.) (2024). *Inteligencia Artificial y Derecho de Daños: Cuestiones Actuales*. Madrid: Dykinson.

Navarro Mendizábal, I. A. (2024). Inteligencia artificial y responsabilidad civil. En J. A. Moreno Martínez y P. J. Femenía López (Coords.), *Inteligencia artificial y Derecho de daños: cuestiones actuales*. Dykinson.

Ortiz Fernández, M. (2024). Los sistemas de inteligencia artificial, ¿productos defectuosos? En J. A. Moreno Martínez y P. J. Femenía López (Coords.), *Inteligencia artificial y Derecho de daños: cuestiones actuales*. Dykinson.

Ortiz Fernández, M. (2026). «Los daños producidos por sistemas inteligentes «defectuosos» en la Directiva europea». *Revista de Derecho Civil*, vol. XIII, núm. 1, pp. 200-235

Oyarzún Vargas, F., "Consideraciones acerca de las acciones de cesación y los daños morales por la pérdida de control de los datos personales. Comentario a la sentencia del TJUE de 4 de septiembre de 2025, asunto C-655/23 (Quirin Privatbank)", *Cuadernos de Derecho Transnacional*, vol. 18, núm. 1, 2026, p. 696.

Pérez García, C. (2026). "Personalidad jurídica e inteligencia artificial", *Universitas*, núm. 48, 2026, pp. 74-97.

Pérez Luño, A. E. (2009). "¿Qué significa juzgar?", *Cuadernos de Filosofía del Derecho*, núm. 32, 2009, pp. 151-170.

Pérez Ureña, A. A. (s.f.). La responsabilidad civil como consecuencia de daños causados por la IA. En V. Magro Servet (Coord.), La responsabilidad civil como consecuencia de daños causados por la Inteligencia Artificial.

Quesada Páez, A. (2026). «La atribución de responsabilidad civil por actos de sistemas autónomos de inteligencia artificial: desafíos para el Derecho privado». InDret, núm. 1, pp. 185-210.

Ramos Martínez, M. F. (2025). "La imputación causal en los daños derivados de la inteligencia artificial", *International Journal of Digital Law*, vol. 6, e604, 2025.

Serra Rodríguez, A., (2026) "El ámbito subjetivo de aplicación de la Directiva (UE) 2024/2853, de 23 de octubre, sobre responsabilidad por los daños causados por productos defectuosos: ¿redefinición de los sujetos protegidos?", *Actualidad Jurídica Iberoamericana*, núm. 24, 2026, pp. 40-60.

Solanes Llano, Á. y Llano Alonso, F. (2026). Ética de la IA y Derechos Humanos. En Actas del VIII Congreso sobre Derechos Humanos e Inteligencia Artificial. Fundación Mainel.

Sotomayor Padilla, P. C., Vaca Nájera, S. O., Vasconez Nolivos, D. C., Moreira Vite, M. D. y Jácome Rocha, A. I. (2026). «Jurisprudencia sobre inteligencia artificial y responsabilidad jurídica». *Ethos Scientific Journal*, vol. 4, núm. 1, pp. 140-160.

Tapia Hermida, A. J. (2024). El seguro de responsabilidad civil profesional de los operadores de sistemas de inteligencia artificial. En J. A. Moreno Martínez y P. J. Femenía López (Coords.), *Inteligencia artificial y Derecho de daños: cuestiones actuales*. Dykinson.

Toribio Carrera, Á. (2026). «El marco regulatorio europeo de la inteligencia artificial y su impacto en el sistema judicial español». *Revista de Estudios Europeos*, núm. 87, pp. 560-580.

4. RECURSOS DE INTERNET

Comunidad de Madrid (2026, 17 de marzo). La Comunidad de Madrid reclama la revisión de los protocolos de valoración de riesgo en el sistema VioGen para mujeres víctimas de violencia. Disponible en: [La Comunidad de Madrid reclama la revisión de los protocolos de valoración de riesgo en el sistema VioGén2 para mujeres víctimas de violencia | Comunidad de Madrid](#) [fecha de consulta: junio 2026].

ECHARRI NICOLÁS, L. y FIERRO RODRÍGUEZ, D., "Crónica de la muerte de la Propuesta de Directiva de daños por inteligencia artificial", *El Español Jurídico (E&J)*, 5 de octubre de 2025. Disponible en: <https://elDerecho.com> [fecha de consulta: marzo 2026].

OLIVA LEÓN, R., "Indemnización de daños causados por sistemas de inteligencia artificial", *Algoritmo Legal*, 15 de diciembre de 2022. Disponible en: <https://www.algoritmolegal.com> [fecha de consulta: marzo 2026].

VILCHES, J., "IA en 'caja negra'", *LegalToday*, 25 de febrero de 2026. Disponible en: <https://www.legaltoday.com> [fecha de consulta: marzo 2026].