Oñati Socio-Legal Series
(ISSN: 2079-5971)

OÑATI INTERNATIONAL INSTITUTE FOR THE SOCIOLOGY OF LAW
OÑATIKO LEGE-SOZIOLOGIAKO NAZIOARTEKO ERAKUNDEA
INSTITUTO INTERNACIONAL DE SOCIOLOGÍA JURÍDICA DE OÑATI

La prevención inteligente del conflicto: inteligencia artificial y la nueva temporalidad del derecho

Oñati Socio-Legal Series Forthcoming
DOI LINK: https://doi.org/10.35295/osls.iisl.2864
ACCEPTED 19 November 2026 | **FIRST-ONLINE PUBLISHED** 25 November 2026

Maria Petronela Popiuc

Profesora en la Facultad de Ciencias Jurídicas
Universidad a Distancia de Madrid (UDIMA) | **Email:** mariapetronela.popiuc@udima.es | **ORCID:** https://orcid.org/0000-0002-6599-0848

Abstracto

La inteligencia artificial (IA) está transformando el momento en el que el derecho interviene ante el conflicto. Los sistemas jurídicos se han configurado tradicionalmente para ofrecer respuestas una vez que la controversia se ha manifestado, pero el análisis de datos y las técnicas de aprendizaje automático permiten identificar indicios tempranos de conflictividad e intervenir antes de que las controversias se intensifiquen. El trabajo sostiene que estas capacidades fundamentan un paradigma sociojurídico emergente de prevención inteligente de los conflictos, denominado derecho anticipativo, que se diferencia tanto del derecho preventivo tradicional como de los modelos privados de gestión del riesgo. A partir de una metodología jurídico-analítica, se examinan sus posibilidades, su encaje normativo europeo y sus desafíos para las garantías del Estado de derecho, y se propone un marco basado en la proporcionalidad, la transparencia, la supervisión humana, la rendición de cuentas y el acceso temprano y efectivo a la justicia.

Palabras clave

Inteligencia artificial, derecho anticipativo, prevención del conflicto, acceso a la justicia, resolución de controversias.

Oñati International Institute for the Sociology of Law
Avenida Universidad, 8 – Apdo. 28
20560 Oñati – Gipuzkoa – Spain
Tel. (+34) 943 783064 / opo@iisj.net / opo.iisj.net

This is an open access work licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International license. (c) The Author(s)

1. Introducción

La inteligencia artificial (IA) ha dejado de ser una herramienta puramente técnica para convertirse en un factor que incide en cómo se organizan las relaciones sociales, se toman decisiones y se gestionan los conflictos. Por eso su llegada al ámbito jurídico plantea algo más que una cuestión de eficiencia y alcanza a la propia función que el derecho desempeña en las sociedades contemporáneas.

Esa función se ha ejercido tradicionalmente *a posteriori*, pues los instrumentos jurídicos de tutela, decisión o composición de controversias, tanto judiciales como extrajudiciales, suelen activarse cuando la situación conflictiva ya es visible. Es precisamente esta premisa la que el desarrollo de la inteligencia artificial y el tratamiento masivo de datos obligan a revisar, en la medida en que permiten reconocer patrones y anticipar escenarios de conflictividad antes de que se conviertan en disputas formalizadas.

Hasta ahora, el debate jurídico sobre inteligencia artificial se ha ocupado sobre todo de su uso en la Administración de Justicia y de la asistencia a jueces y abogados. Este artículo parte, en cambio, de una pregunta distinta, que no es si la IA puede hacer más eficiente el sistema jurídico, sino si su capacidad anticipativa puede desplazar el momento en que el derecho interviene frente al conflicto, y bajo esa mirada la IA deja de ser una simple herramienta auxiliar para convertirse en un instrumento orientado a la prevención inteligente del conflicto.

Una prevención de este tipo puede aportar beneficios notables, como la detección temprana de riesgos, un mejor acceso a información jurídica en las fases iniciales y el fomento de respuestas colaborativas antes de que la controversia escale, aunque también puede intensificar la vigilancia, reforzar desigualdades o trasladar decisiones relevantes a sistemas opacos. De ahí que su análisis exija valorar no solo su eficacia práctica, sino también sus condiciones de legitimidad jurídica e institucional.

Con este punto de partida, el artículo se pregunta si la IA puede contribuir a configurar un modelo anticipativo de intervención jurídica sin comprometer los principios del Estado de derecho, y sostiene que su principal aportación no está en automatizar la decisión jurídica, sino en permitir que el derecho intervenga antes y con mayor capacidad de evitar la escalada de las controversias.

Para responder a esa pregunta se sigue una metodología de dogmática jurídica cualitativa, que combina el análisis doctrinal sobre la función del derecho frente al conflicto con el examen del marco normativo europeo en materia de inteligencia artificial, protección de datos, servicios digitales y trabajo en plataformas. A ello se suma el estudio de aplicaciones actuales de la IA en ámbitos especialmente expuestos a la conflictividad, como las relaciones de consumo, las plataformas digitales y la Administración pública. Se trata, por tanto, de un trabajo de naturaleza conceptual, en el que los ejemplos de plataformas cumplen una función ilustrativa y se basan en la información que las propias empresas hacen pública, sin pretender ofrecer evidencia empírica de su funcionamiento.

Además de esta introducción, el artículo se organiza en cinco apartados. El segundo examina la función tradicional del derecho frente al conflicto y los límites del paradigma reactivo en las sociedades digitales, lo que permite al tercero abordar la IA como tecnología de anticipación, delimitar el derecho anticipativo frente al derecho preventivo y la gestión del riesgo y contrastar con esos criterios algunas aplicaciones actuales. Sobre esa base,

el cuarto desarrolla la propuesta, con sus principios, el marco normativo europeo en que se inserta y su integración en una arquitectura de justicia temprana, mientras que el quinto analiza sus riesgos y límites y las conclusiones precisan las condiciones de su aplicación.

2. La función tradicional del derecho frente al conflicto

2.1. El paradigma reactivo del derecho

Antes de examinar la capacidad anticipatoria de la inteligencia artificial conviene detenerse en el punto de partida, que no es otro que una relación entre derecho y conflicto construida históricamente sobre una estructura de intervención posterior. A ello se añade que, desde la perspectiva sociojurídica, el conflicto no entra de forma inmediata en el sistema jurídico como disputa formalizada, sino que atraviesa antes un proceso en el que una experiencia lesiva se percibe como tal, se atribuye a otro sujeto y finalmente se expresa como reclamación.

Felstiner, Abel y Sarat conceptualizaron este proceso mediante la conocida secuencia *naming, blaming, claiming*, con la que explican cómo determinadas experiencias de daño llegan, o no, a convertirse en disputas jurídicamente reconocibles (Felstiner *et al.* 1981, 631–633). En una línea próxima, Miller y Sarat sostienen que el estudio de la resolución institucional no debería partir de la disputa ya formada, sino del proceso previo por el cual los agravios se transforman en quejas y controversias (Miller y Sarat 1981, 525). De ahí que, incluso desde una perspectiva estrictamente jurídica, los mecanismos de solución de controversias presupongan de ordinario la existencia previa de una disputa reconocible (Schreuer 2008, 959), y que el derecho no suela intervenir en la fase inicial de formación del conflicto, sino cuando este ha adquirido una forma lo bastante determinada como para traducirse al lenguaje institucional del sistema jurídico.

El proceso judicial es la expresión más depurada de esta racionalidad reactiva, puesto que la jurisdicción presupone una controversia formalizada entre partes y la solicitud de tutela ante un órgano dotado de autoridad para resolverla. La teoría jurídica anglosajona denomina *adjudication* (adjudicación) a esta forma de ordenación social, en la que un tercero imparcial resuelve una controversia entre partes con arreglo a normas o criterios preestablecidos y que comprende, por tanto, tanto la función jurisdiccional como el arbitraje. Fuller señaló que su rasgo característico es ofrecer a las partes una forma particular de participación en la decisión, consistente en presentar pruebas y argumentos razonados ante un tercero imparcial (Fuller 1978, 353-409). Para que el conflicto pueda tratarse procesalmente debe reconstruirse, además, mediante categorías jurídicas (hechos, fundamentos normativos, legitimación, carga probatoria, decisión), y esa traducción confirma que la intervención judicial llega cuando el conflicto ya se ha consolidado como litigio.

Tampoco los mecanismos extrajudiciales escapan del todo a esta estructura temporal. La mediación, el arbitraje, la conciliación y la negociación asistida se han presentado como respuestas más flexibles y menos adversariales que el proceso judicial (Sander 1976, 321, Menkel-Meadow 1991, 1-46), pero también ellos operan, por regla general, sobre una

controversia ya existente. La mediación presupone un desacuerdo susceptible de diálogo, el arbitraje una disputa sometida a un tercero y la negociación, incluso en sus fases más tempranas, suele partir de intereses contrapuestos ya identificados (Moore 2014, 23, Blackaby *et al.* 2022). Han diversificado así las técnicas de tratamiento del conflicto e introducido una racionalidad más cooperativa, sin llegar a superar la matriz reactiva de la intervención jurídica tradicional.

Ello no significa que el derecho carezca de dimensiones preventivas, pues numerosos sectores normativos incorporan reglas dirigidas a evitar daños o a imponer deberes de diligencia, como ocurre en la prevención de riesgos laborales, en los deberes de información en materia de consumo o en los sistemas internos de alerta ante infracciones normativas (Ley 31/1995, art. 2; Real Decreto Legislativo 1/2007, art. 12; Ley 7/2022, art. 1; Ley 2/2023, preámbulo y tít. II). Se trata, sin embargo, de una prevención normativa y estructural, que reduce categorías generales de riesgo sin llegar a identificar de forma individualizada la formación de un conflicto concreto, y una cosa es que el derecho establezca normas preventivas generales y otra bien distinta que disponga de mecanismos capaces de detectar tempranamente un conflicto específico y de activar respuestas proporcionadas antes de su escalada. Por eso, con frecuencia, el derecho llega cuando el daño ya se ha producido, la confianza se ha deteriorado o las posiciones se han polarizado.

Esa llegada tardía no es solo temporal, sino también epistemológica, puesto que el derecho solo puede tratar plenamente lo que reconoce con sus propias categorías y una situación conflictiva debe traducirse a conceptos como incumplimiento, daño o vulneración de derechos para ingresar en el sistema jurídico. Bourdieu explicó que entrar en el campo jurídico implica aceptar sus reglas específicas, hasta el punto de que solo permanece en el caso aquello argumentable desde la pertinencia legal (Bourdieu 1987, 831-832). En términos de Luhmann, el derecho solo opera plenamente cuando el conflicto se ha traducido a su código de lo lícito y lo ilícito (Luhmann 2004, 93-94), de forma que antes de esa traducción apenas dispone de instrumentos para percibirlo. No es extraño, por ello, que muchas controversias sean en realidad la fase visible de procesos previos de deterioro relacional, desinformación o asimetría de poder que no fueron gestionados a tiempo.

El paradigma reactivo presenta, en consecuencia, una ambivalencia estructural, porque resulta imprescindible para garantizar seguridad jurídica y tutela judicial efectiva, pero muestra límites evidentes frente a conflictos complejos, masivos o tecnológicamente mediados, en los que la intervención posterior puede resultar insuficiente para evitar la escalada o la consolidación de desigualdades. El problema no está, por tanto, en la existencia de un derecho reactivo, sino en que ocupe en exclusiva el centro de la arquitectura jurídica en sociedades marcadas por la aceleración de las relaciones sociales y la creciente complejidad de la conflictividad.

2.2. Los límites del modelo tradicional en las sociedades digitales

Esos límites se han agudizado con la digitalización de las relaciones sociales, a medida que la expansión de las plataformas digitales, la automatización de procesos y la adopción de decisiones tecnológicamente mediadas transforman las condiciones en que los conflictos se generan y escalan. Las controversias contemporáneas ya no siempre se

presentan como disputas individualizadas entre sujetos en condiciones simétricas, sino que surgen con frecuencia en entornos dinámicos, masivos y técnicamente complejos, donde la capacidad de reacción de las personas afectadas está limitada (Micklitz 2021, 298, Ortiz Laverde y Ortiz Cabrera 2025, 1).

Esa limitación se agrava por la velocidad a la que se producen muchos conflictos digitales, difícilmente compatible con los tiempos ordinarios de respuesta jurídica, pues la suspensión de una cuenta, la denegación automatizada de un servicio o una decisión basada en datos pueden tener efectos inmediatos, y cuando llega la intervención posterior el daño ya se ha producido y resulta difícilmente reversible. A la rapidez se une el carácter masivo y repetitivo de esta conflictividad, porque un mismo diseño contractual o un mismo sistema automatizado puede afectar simultáneamente a un número muy elevado de personas, mientras que el modelo reactivo tiende a procesar estos problemas mediante reclamaciones individuales y encuentra, por ello, dificultades para identificar a tiempo los patrones estructurales de conflictividad (Chauhan 2020, 285, Lagioia *et al.* 2022, 481).

A todo ello se suma que las sociedades digitales incrementan las asimetrías informativas y económicas entre las partes, puesto que una de ellas controla la infraestructura tecnológica y conoce los criterios empleados para decidir, mientras que la otra puede desconocer qué información se ha usado o qué vía de reclamación le corresponde. Esta asimetría compromete el acceso efectivo a la justicia (Larsson 2017, 67, Bloch-Wehba 2020, 1265) y se ve agravada por una opacidad estructural que no deriva solo de la complejidad técnica, sino también del modo en que las organizaciones diseñan y justifican sus decisiones. Cuando la causa de la afectación no es visible, la respuesta jurídica reactiva se debilita, porque el conflicto existe pero permanece parcialmente oculto para quien lo padece (Battaglini y Rasmussen 2019, 331, Castets-Renard 2019, 91).

Rapidez, masividad, asimetría y opacidad ponen así de manifiesto las dificultades del modelo tradicional ante unas formas de conflictividad en las que el conflicto no comienza necesariamente cuando se presenta una demanda, sino que con frecuencia se gesta antes, mediante la acumulación de señales débiles, incumplimientos leves o prácticas organizativas que incrementan progresivamente el riesgo de controversia. Si el derecho interviene solo en la fase final, corre el riesgo de limitarse a gestionar las consecuencias de un conflicto que podría haberse canalizado antes de su escalada.

De este diagnóstico no se deriva que deban sustituirse los mecanismos tradicionales de tutela, sino que conviene complementarlos con formas de detección temprana, orientación jurídica anticipada y gestión previa de los riesgos. Esa necesidad se comprende mejor si se tiene en cuenta que el derecho no llega tarde por una simple ineficiencia operativa, sino por su propio diseño institucional, cuyas categorías y procedimientos se construyeron para actuar cuando el conflicto ya es visible y procesable por las instituciones.

La cuestión de fondo es, entonces, si el derecho puede incorporar una función anticipativa sin perder las garantías que sostienen su legitimidad, lo que no consiste en atribuir a la tecnología una capacidad neutral de solución, sino en examinar bajo qué condiciones puede servir para identificar riesgos y favorecer respuestas tempranas compatibles con los derechos fundamentales. A esa cuestión se dedican los apartados siguientes.

3. La inteligencia artificial y el nacimiento de un derecho anticipativo

3.1. La capacidad anticipativa de la inteligencia artificial

Frente a ese diagnóstico, cabe preguntarse qué cambia realmente con la llegada de la IA. A diferencia del modelo tradicional, centrado en responder a controversias ya formalizadas, los sistemas inteligentes permiten identificar patrones y anticipar escenarios de conflictividad antes de que se consoliden como litigios o disputas institucionalizadas, y en ello reside su verdadero interés, que no es tanto automatizar tareas jurídicas como desplazar parcialmente la intervención jurídica hacia una fase anterior del conflicto (Schoinas y Qadir 2017, 3835, González Fernández 2024, 51, Katsh *et al.* 2025, 79).

Esta capacidad anticipativa procede de la posibilidad de procesar grandes volúmenes de datos, reconocer regularidades e identificar anomalías a partir de la información acumulada, gracias a lo cual el conflicto deja de observarse como un hecho consumado y pasa a entenderse como un proceso gradual, precedido por señales débiles como reclamaciones reiteradas, patrones de insatisfacción o prácticas organizativas que incrementan el riesgo de controversia (Corbett 2016, 5-6). La aportación de la IA consiste precisamente en pasar del caso aislado al patrón de conflictividad, porque una reclamación individual puede parecer un incidente singular, mientras que una pluralidad de reclamaciones semejantes puede revelar un problema estructural, en línea con la lógica del *dispute systems design*, que favorece procesos de aprendizaje y prevención a escala sistémica (Rabinovich-Einy y Katsh 2012, 156-157).

Esa capacidad predictiva no equivale, sin embargo, a certeza jurídica. Los modelos de aprendizaje automático permiten formular predicciones sobre nuevos supuestos sin que ello suponga una comprensión jurídica del conflicto, y en ocasiones producen resultados difíciles de explicar incluso para quienes los entrenan (Mittelstadt *et al.* 2016, 3–6), por lo que la IA puede detectar riesgos, pero no determinar por sí sola su relevancia normativa. Decidir si una situación exige revisión humana, mediación o tutela judicial sigue requiriendo criterios jurídicos y control institucional, y en este sentido las directrices europeas sobre IA confiable insisten en que estos sistemas deben apoyar la decisión humana y permitir formas adecuadas de supervisión (High-Level Expert Group on Artificial Intelligence 2019, 16–17).

Por ello, la IA debe concebirse como herramienta auxiliar de detección y apoyo a la decisión, y no como instancia autónoma de prevención o resolución del conflicto.

Entendida de este modo, la IA puede contribuir a configurar una fase previa de detección, diagnóstico y orientación temprana que no sustituye a la negociación, la mediación o el proceso judicial, sino que puede precederlos e incluso, en algunos casos, evitar que el conflicto llegue a ellos, sobre todo cuando permite intervenir antes de que la controversia se consolide plenamente (Rabinovich-Einy y Katsh 2012, 152, 194-195). El derecho anticipativo no implica, por tanto, abandonar la función reactiva del derecho, sino complementarla con instrumentos capaces de corregir las dinámicas que generan controversias antes de que estas se consoliden.

Esa anticipación exige, con todo, cautela, porque prevenir conflictos no puede significar vigilar indiscriminadamente las relaciones sociales ni perfilar preventivamente a personas o colectivos sobre la base de predicciones opacas (Yeung 2018). Los riesgos de esa deriva se examinan más adelante, pero antes conviene precisar qué debe entenderse por derecho anticipativo y en qué se diferencia de figuras afines.

3.2. Delimitación conceptual: derecho anticipativo, derecho preventivo y gestión del riesgo

La expresión “derecho anticipativo” requiere una delimitación precisa, puesto que el ordenamiento conoce desde antiguo figuras preventivas con las que podría confundirse. A efectos de este trabajo, se entiende por derecho anticipativo la función jurídica que, apoyándose en el tratamiento de datos, identifica señales tempranas de un conflicto concreto, referido a personas o relaciones determinadas, y activa respuestas proporcionadas y revisables (información, revisión humana, derivación a un mecanismo de solución) antes de que ese conflicto se formalice como reclamación o litigio. Esa función se ejerce, además, dentro de un marco de garantías que permite a la persona afectada conocerla y oponerse a ella.

Así entendido, el derecho anticipativo se separa ante todo del derecho preventivo clásico, cuyas normas de prevención de riesgos laborales, deberes de información en materia de consumo o canales internos de denuncia operan *ex ante*, pero de forma general y abstracta, reduciendo categorías de riesgo sin identificar la gestación de un conflicto singular. Tampoco se confunde con las técnicas procesales de anticipación, como la prueba anticipada o las medidas cautelares, que presuponen ya una controversia definida.

Lo que lo distingue es la individualización temprana, porque actúa sobre un conflicto concreto cuando todavía no ha sido reclamado por sus protagonistas o, en la secuencia de Felstiner, Abel y Sarat, mientras el agravio aún se está nombrando o atribuyendo. No es casual, por ello, que Katsh, Schwartz-Maor y Rabinovich-Einy hayan propuesto añadir a esa secuencia una fase de prevención que la IA hace posible (Katsh *et al.* 2025, 79).

Más próxima, y por ello más necesitada de deslinde, es la gestión algorítmica del riesgo que las empresas desarrollan en su propio interés, con la que el derecho anticipativo comparte la técnica (detección de patrones, puntuación de riesgos, alertas), aunque no el propósito. La gestión del riesgo protege ante todo el interés de quien controla el sistema, que busca evitar pérdidas, fraudes o reclamaciones, y por eso suele traducirse en medidas unilaterales contra quien aparece como sospechoso, ya se trate de bloquear una operación, desactivar una cuenta o denegar un servicio. El derecho anticipativo mira, en cambio, a la persona que puede resultar lesionada y a la escalada del conflicto entre las partes, y responde informando, protegiendo o abriendo un cauce de solución. De esa diferencia de finalidad se deriva la que resulta decisiva, pues la intervención anticipativa solo merece ese nombre si es conocida, motivada y revisable por una instancia humana e independiente.

Individualización temprana, orientación a la protección de derechos y sujeción a garantías no describen tipos puros, sino un gradiente, pero sirven para examinar las aplicaciones actuales de la IA y distinguir cuáles se aproximan al modelo propuesto y cuáles constituyen gestión privada del riesgo con efectos preventivos meramente reflejos.

3.3. Aplicaciones actuales de la inteligencia artificial a la luz de esos criterios

Para aplicar estos criterios conviene empezar por distinguir la IA preventiva de la mera digitalización de los canales de reclamación, porque no todo formulario en línea o sistema de resolución de controversias en línea (ODR) constituye una aplicación de IA propiamente dicha. Vilalta Nicuesa y Gili Saldaña diferencian, en este sentido, los mecanismos de ADR/ODR asistidos por IA según su grado de autonomía y riesgo (Vilalta Nicuesa y Gili Saldaña 2025, 1-18). Incluso cuando incorporan IA, además, las plataformas ODR intervienen cuando el conflicto ya ha llegado a ellas, y aunque pueden acortar y abaratar su tratamiento, pertenecen en rigor al momento reactivo. La especificidad de la IA preventiva reside, en cambio, en su capacidad para clasificar información, detectar patrones y recomendar respuestas antes de que exista una reclamación, como muestra el uso que algunas administraciones públicas hacen ya de ella para predecir riesgos e identificar supuestos que requieren inspección (Cerrillo i Martínez 2021, 3). Partiendo de una concepción de la IA en la justicia alternativa como herramienta de apoyo y no de sustitución (Gonzalo Quiroga 2024, 147-148), se examinan a continuación varios ejemplos, ordenados de menor a mayor proximidad al modelo propuesto.

El primero procede de los pagos digitales, donde plataformas como PayPal emplean aprendizaje automático para detectar operaciones sospechosas, en línea con una tendencia más amplia de minería de datos aplicada a la detección de fraude financiero (Bolton y Hand 2002, 235–236, Ngai *et al.* 2011, 559–560, PayPal 2024, Das 2025). Estos sistemas evitan muchos contracargos y conflictos entre usuarios, comerciantes y proveedores, pero lo hacen como efecto reflejo de su verdadero objetivo, que es proteger a la plataforma y a sus clientes frente al fraude, de ahí que su respuesta típica sea unilateral (retener o bloquear la operación) y que su lugar esté más cerca de la gestión privada del riesgo que del derecho anticipativo.

Algo parecido ocurre con Amazon, que declara usar IA para detectar reseñas falsas antes de su publicación y reforzar los controles frente a productos falsificados (Amazon 2024, Mehta 2025). Los sistemas de valoración en línea son centrales para la confianza en los mercados electrónicos, pero pueden manipularse mediante reseñas falsas que alteran la decisión del consumidor (Dellarocas 2003, 1407–1408, Mayzlin *et al.* 2014, 2421–2422, Luca y Zervas 2016, 3412–3413), y la IA actúa aquí sobre el diseño del mercado antes de que el consumidor decida con información manipulada. Su efecto preventivo es, sin embargo, general, porque no identifica un conflicto concreto entre personas determinadas y se aproxima más a la prevención estructural que a la anticipación individualizada; como en el caso de PayPal, lo que cambia es la escala de la vigilancia y no su naturaleza, que sigue siendo la de los controles antifraude tradicionales.

Airbnb ofrece un tercer ejemplo, con sistemas de cribado que identifican el riesgo de fiestas no autorizadas o de daños a la propiedad y que bloquean o revisan reservas antes de que se produzca el daño (Airbnb 2026). En este caso el sistema sí individualiza, porque opera sobre una reserva y una persona concretas, pero responde restringiendo el acceso al servicio con criterios no siempre transparentes para el usuario, una ambivalencia característica de las plataformas de economía colaborativa (Calo y Rosenblat 2017, 1628–1629). Por eso el caso permite ver con claridad dónde se sitúa la frontera, puesto que

existe anticipación, pero al servicio de la plataforma y del anfitrión y sin garantías que permitan a la persona afectada conocer y discutir la decisión.

La moderación de contenidos en plataformas como YouTube, Facebook, Instagram o TikTok se aleja más de las figuras meramente preventivas. Los conflictos masivos por discurso de odio, desinformación o derechos de autor han convertido la moderación algorítmica en pieza central de la gobernanza de plataformas, en la medida en que permite gestionar a gran escala contenidos de riesgo (Gorwa *et al.* 2020, 1-3, Meta 2023, YouTube 2026). No estamos, sin embargo, ante una versión ampliada de controles previos, sino ante un fenómeno nuevo, en el que la plataforma decide, a gran escala y antes de cualquier reclamación, sobre la colisión entre la libertad de expresión de quien publica y los derechos de terceros. Esa misma lógica explica que la moderación genere a su vez nuevas controversias cuando la retirada de contenidos o la suspensión de cuentas se produce sin revisión humana efectiva (Gillespie 2018, 5-6), y no sorprende, por ello, que el Reglamento de Servicios Digitales haya respondido a este fenómeno con garantías propias de una lógica anticipativa.

Algo similar sucede en las plataformas de movilidad como Uber, cuyos sistemas de detección de fraude o incumplimiento pueden prevenir daños, pero también generar conflictos sensibles cuando dan lugar a desactivaciones de cuenta difíciles de impugnar (Uber 2026). Uber ejerce formas intensas de control mediante reputación e incentivos algorítmicos que condicionan el acceso al trabajo (Rosenblat y Stark 2016, 3758-3759), y determinadas métricas pueden situar a los conductores en riesgo de desactivación, lo que evidencia la asimetría de poder entre plataforma y prestador (Calo y Rosenblat 2017, 1661-1663). El ejemplo muestra, así, que la anticipación puede ser ella misma fuente de conflicto, y explica que el legislador europeo haya sometido la gestión algorítmica del trabajo en plataformas a deberes de transparencia, supervisión y revisión humanas.

Un ejemplo más próximo al modelo propuesto puede encontrarse en la gestión de las solicitudes de retirada de resultados de búsqueda en ejercicio del derecho de supresión. Desde la sentencia *Google Spain* (TJUE 2014) y, después, al amparo del artículo 17 del Reglamento General de Protección de Datos, Google resuelve un elevado volumen de solicitudes en las que debe ponderar el derecho a la protección de datos de una persona física concreta y el derecho a la información. Según la propia empresa, en ese proceso intervienen tanto la automatización como la revisión por su personal (Google 2026a, 2026b). La experiencia acumulada permitiría identificar información que muy probablemente dará lugar a solicitudes estimables, como las noticias sobre acusaciones de las que la persona fue absuelta, y, sin esperar a su reclamación, informar a la persona afectada y ofrecerle la retirada. Este escenario, hoy hipotético, reúne los tres criterios expuestos, porque individualiza un conflicto antes de que sea reclamado, se orienta a proteger los derechos de la persona potencialmente afectada y deja en sus manos la decisión, con posibilidad de revisión por la autoridad de protección de datos. Exigiría, con todo, resolver a su vez cuestiones de minimización de datos y no presuponer la voluntad de la persona interesada.

Vistas en conjunto, estas aplicaciones muestran que la prevención inteligente del conflicto no es neutral, puesto que puede facilitar la detección temprana de riesgos, pero también reforzar asimetrías de poder o imponer soluciones automatizadas opacas y difíciles de impugnar (Mittelstadt *et al.* 2016, 6–10, Yeung 2018). Como la mayoría de las

prácticas actuales se sitúan del lado de la gestión privada del riesgo, el derecho anticipativo no describe lo que las plataformas ya hacen, sino que ofrece un modelo normativo con el que evaluarlo y corregirlo. La cuestión central deja de ser, así, si la IA permite anticipar controversias, para pasar a ser bajo qué condiciones jurídicas debe hacerlo, lo que exige una teoría capaz de definir sus finalidades, principios y límites.

4. Hacia una teoría sociojurídica del derecho anticipativo

Los ejemplos analizados muestran que la inteligencia artificial ya se emplea, de facto, para anticipar riesgos de conflictividad, pero también que esa capacidad necesita un marco jurídico propio que la oriente y la limite. En ese sentido, la IA permite plantear el tránsito desde un derecho predominantemente reactivo hacia un derecho anticipativo, cuya aportación no está en automatizar la decisión jurídica ni en sustituir al juez, al mediador o a las partes por sistemas algorítmicos, sino en reorganizar el ciclo jurídico del conflicto incorporando una fase previa de detección, diagnóstico y orientación. En esa fase la IA actúa como apoyo para identificar patrones y facilitar respuestas tempranas, mientras que la valoración jurídica de los riesgos permanece sometida a criterios humanos e institucionales. Así lo entiende la literatura sobre IA y resolución de conflictos, que coincide en atribuirle una función asistencial antes que decisoria (Solar Cayón 2020, González Fernández 2024, 52), y en esa misma línea se sitúan los estudios que analizan los límites de esa asistencia en la toma de decisiones judiciales (Ester Sánchez 2025, 318).

Desde una perspectiva sociojurídica, esta propuesta parte de la premisa de que los conflictos no nacen jurídicamente formados, sino que se construyen progresivamente a través de experiencias de daño, atribuciones de responsabilidad, asimetrías de poder o ausencia de canales eficaces de respuesta. Mientras el derecho tradicional interviene cuando ese proceso ya ha cristalizado en litigio, el derecho anticipativo se sitúa en una fase anterior, cuando todavía es posible informar, revisar, corregir o derivar el desacuerdo hacia mecanismos adecuados de solución. La premisa recoge la idea de Felstiner, Abel y Sarat de que la disputa es el resultado de un proceso de transformación (Felstiner *et al.* 1981, 631–633) y conecta con la concepción del derecho como sistema de comunicación orientado al entendimiento social y no solo a la decisión (Medina Morales 2023, 95–116), así como con la tradición de prácticas autocompositivas que, en distintas culturas, tratan el conflicto al margen del juicio (Moreno Buján 2021).

La novedad de este planteamiento no consiste, por tanto, en eliminar el conflicto, aspiración incompatible con una sociedad plural, sino en mejorar la calidad temporal de la respuesta jurídica. Allí donde el derecho reactivo actúa cuando el daño ya se ha producido o las posiciones se han polarizado, el derecho anticipativo introduce una capa previa de prevención inteligente que no busca reducir artificialmente la litigiosidad, sino que el derecho llegue antes, con menor intensidad y con mayor capacidad de evitar escaladas innecesarias.

4.1. Los principios del derecho anticipativo

La anticipación, por sí sola, no es un valor jurídico absoluto, porque puede servir para evitar daños, pero también para vigilar, perfilar o condicionar conductas antes de que exista una controversia real. Por eso el derecho anticipativo solo puede ser legítimo si se apoya en una arquitectura de principios capaz de limitar el poder preventivo que la IA confiere a instituciones, empresas y plataformas.

El primero de esos principios es la propia anticipación, entendida no como predicción cerrada del futuro, sino como capacidad de reconocer señales tempranas de conflictividad (quejas reiteradas, incumplimientos leves, prácticas contractuales opacas) antes de que un conflicto adquiera apariencia litigiosa. Anticipar no significa, en este sentido, actuar sobre meras sospechas, sino ampliar la capacidad de diagnóstico del sistema jurídico para evitar un daño mayor.

Esa intervención temprana debe estar limitada por la proporcionalidad, de manera que una alerta algorítmica pueda justificar información adicional o revisión humana, pero no necesariamente una sanción o una restricción de derechos. Cuanto mayor sea el impacto de la medida anticipativa, mayor debe ser su justificación, porque el riesgo probable no puede tratarse como daño consumado ni la predicción ocupar el lugar de la prueba.

La proporcionalidad exige, a su vez, una supervisión humana efectiva. La IA puede clasificar riesgos y recomendar respuestas, pero no debe sustituir la valoración jurídica, y la supervisión ha de permitir comprender la alerta, revisar sus fundamentos y apartarse de la recomendación cuando resulte necesario, sobre todo cuando la intervención afecta al acceso a un servicio o a la admisión de una reclamación (High-Level Expert Group on Artificial Intelligence 2019, 16–18, Unesco 2021, 23–24, Parlamento Europeo y Consejo 2024a, arts. 14 y 26).

A ella se suma la transparencia, en virtud de la cual la persona afectada por una alerta o medida preventiva debe poder conocer, en términos comprensibles, por qué ha sido objeto de esa intervención, sin que ello obligue a revelar todos los elementos técnicos del sistema. Sin transparencia, la prevención puede acabar generando la misma opacidad que pretendía evitar (Pasquale 2015, 19, Burrell 2016, 1–12, Wachter *et al.* 2017, 76–99, Kaminski 2020, 121–138), y es también la transparencia la que sostiene la confianza, porque la prevención inteligente solo resulta aceptable si las personas perciben que no se utiliza para cerrar reclamaciones o proteger exclusivamente los intereses de quien controla el sistema, lo que exige información clara, trazabilidad y garantías verificables.

Por último, el derecho anticipativo debe orientarse al acceso temprano a la justicia, entendiendo que prevenir el conflicto no significa impedir o desactivar la reclamación, sino facilitar que la persona afectada acceda antes a información, orientación y mecanismos adecuados de solución, lo que amplía la concepción clásica de acceso a la justicia más allá del juzgado (Cappelletti y Garth 1978, 181, Susskind 2019, 253, Hodges 2019).

Son estos principios los que permiten distinguir el derecho anticipativo de una mera gestión algorítmica del riesgo, porque la anticipación sin proporcionalidad puede convertirse en control excesivo, la prevención sin transparencia puede generar indefensión y la eficiencia sin acceso temprano a la justicia puede invisibilizar conflictos legítimos. Por ello, el derecho anticipativo debe construirse como una teoría jurídica de la prevención con garantías y no como una técnica de optimización institucional.

4.2. El marco normativo europeo del derecho anticipativo

Aunque ninguna norma europea formula el derecho anticipativo como tal, el derecho de la Unión ofrece ya piezas que permiten articularlo y, al mismo tiempo, limitarlo, entre las que destacan cuatro instrumentos.

El punto de partida es el Reglamento General de Protección de Datos, puesto que toda detección temprana de conflictos exige tratar datos personales y ese tratamiento queda sujeto a los principios de limitación de la finalidad y de minimización, a la protección de datos desde el diseño y, cuando entraña un alto riesgo, a una evaluación de impacto previa. El artículo 22 reconoce, además, el derecho a no ser objeto de decisiones basadas únicamente en el tratamiento automatizado, incluida la elaboración de perfiles, que produzcan efectos jurídicos o afecten significativamente de modo similar. En los supuestos en que esas decisiones se admiten, el Reglamento exige garantías como la intervención humana, la posibilidad de expresar el propio punto de vista y la de impugnar la decisión (Parlamento Europeo y Consejo 2016, arts. 5.1.b y c, 22, 25 y 35). Son estas reglas las que impiden que la anticipación se convierta en vigilancia generalizada y que la alerta algorítmica se transforme sin más en decisión.

A esa base se añade el Reglamento (UE) 2024/1689, de inteligencia artificial, cuyo enfoque basado en el riesgo resulta especialmente útil para graduar la intervención anticipativa. En el extremo, prohíbe prácticas que encarnan la forma más grave de anticipación ilegítima, como la puntuación social o la evaluación del riesgo de que una persona física cometa un delito basada únicamente en la elaboración de perfiles. Califica, además, de alto riesgo, entre otros, los sistemas que condicionan el acceso a servicios privados y públicos esenciales, los empleados en la gestión del trabajo y los destinados a asistir a las autoridades judiciales o a utilizarse de manera similar en la resolución alternativa de litigios. Para estos últimos impone supervisión humana efectiva, obligaciones específicas a quienes los despliegan, una evaluación de impacto en los derechos fundamentales en determinados supuestos y el derecho de la persona afectada a obtener explicaciones sobre la decisión individual (Parlamento Europeo y Consejo 2024a, arts. 5.1.c y d, 6, 14, 26, 27 y 86, y anexo III, puntos 4, 5 y 8.a).

Es, sin embargo, el Reglamento (UE) 2022/2065, de servicios digitales, el que contiene probablemente la arquitectura más cercana a la lógica anticipativa. En él se obliga a los prestadores de servicios de alojamiento de datos a motivar las restricciones que imponen a los usuarios, se exige a las plataformas en línea un sistema interno de gestión de reclamaciones y se reconoce el acceso a órganos certificados de resolución extrajudicial de litigios. A las plataformas y motores de búsqueda de muy gran tamaño les impone, además, evaluar y mitigar los riesgos sistémicos que generan sus servicios, lo que supone identificar *ex ante* patrones de daño y corregir el diseño que los produce (Parlamento Europeo y Consejo 2022, arts. 17, 20, 21, 34 y 35). Este esquema, que combina detección del riesgo, corrección estructural, motivación individual y vías tempranas de solución, coincide en buena medida con el ciclo de detección, diagnóstico y respuesta que aquí se propone.

Cierra este conjunto la Directiva (UE) 2024/2831, relativa a la mejora de las condiciones laborales en el trabajo en plataformas, que responde directamente a situaciones como las descritas en el caso de Uber. Para ello exige transparencia sobre los sistemas

automatizados de seguimiento y de toma de decisiones, su supervisión humana y el derecho de la persona trabajadora a obtener una explicación y a solicitar la revisión de las decisiones adoptadas o respaldadas por ellos (Parlamento Europeo y Consejo 2024b, arts. 9, 10 y 11).

Todo este conjunto normativo proporciona al derecho anticipativo sus límites externos y algunos de sus instrumentos, pero su carácter fragmentario explica que siga siendo necesaria una construcción teórica que dé coherencia a esas piezas desde la perspectiva del conflicto y del acceso a la justicia.

4.3. La inteligencia artificial como infraestructura de gobernanza preventiva

En el marco del derecho anticipativo, la IA debe entenderse como una infraestructura de gobernanza preventiva, cuya función no consiste únicamente en resolver casos concretos, sino en producir información sobre cómo se generan, repiten y escalan los conflictos. En ello reside su principal aportación sociojurídica, porque es lo que permite pasar de la gestión aislada de reclamaciones a la identificación de patrones de conflictividad.

Esa infraestructura opera en tres planos. El primero es la detección, en la medida en que los sistemas inteligentes pueden identificar señales que no resultan visibles mediante el análisis individual de cada reclamación (productos con quejas recurrentes, cláusulas que generan confusión, procedimientos administrativos que producen recursos repetidos) y advertir que lo que parecía un incidente aislado puede ser la manifestación de una dinámica más amplia (Surden 2014, 89–94, Coglianese y Lehr 2017, 1160–1165, Katsh y Rabinovich-Einy 2017).

Detectar, sin embargo, no basta, y el valor jurídico de la IA crece en el segundo plano, el del diagnóstico, cuando permite relacionar prácticas, diseños contractuales o procesos organizativos con determinados patrones de conflicto, como ocurre cuando una plataforma advierte que ciertas suspensiones de cuenta generan impugnaciones recurrentes o una Administración comprueba que un procedimiento produce recursos por falta de claridad. Se pasa así de la reacción frente al caso individual a la corrección de la causa estructural.

Sobre esa base se construye el tercer plano, el de la respuesta temprana, que solo tiene sentido si permite activar actuaciones proporcionadas (información personalizada, revisión humana, derivación a mediación, rediseño de una interfaz) que lleguen antes y respeten las garantías de las personas afectadas. La doctrina procesalista sobre los llamados *smart ODR* distingue, en este sentido, entre los usos asistenciales de la IA y sus posibles usos decisorios (González Fernández 2022, Marcos Francisco 2023, 1–41), y es en los primeros donde la respuesta temprana encuentra su espacio natural (Gonzalo Quiroga 2024, 147–152).

Desde esta perspectiva, la inteligencia artificial no ocupa el centro del derecho anticipativo, sino que actúa como su soporte técnico, mientras que el centro sigue siendo jurídico y lo integran la finalidad legítima, la proporcionalidad, la transparencia, la supervisión humana y el acceso temprano a la justicia. Esta concepción permite, además, evitar dos desviaciones opuestas, la de reducir el derecho anticipativo a una promesa de eficiencia, como si prevenir el conflicto se midiera solo por la reducción de litigios cuando

también puede disminuirse la litigiosidad mediante obstáculos de acceso u opacidad, y la de rechazar toda anticipación por identificarla automáticamente con control algorítmico.

La cuestión no es, por tanto, si el derecho puede anticiparse, sino bajo qué condiciones puede hacerlo legítimamente. La respuesta pasa por reconocer que la IA solo funciona como infraestructura de gobernanza preventiva cuando se orienta a detectar problemas antes invisibles y a abrir vías tempranas de solución, mientras que pierde esa legitimidad en cuanto se emplea para clasificar personas, bloquear reclamaciones o cerrar prematuramente el conflicto.

4.4. Derecho anticipativo y arquitectura de justicia temprana

De ahí que el derecho anticipativo no pueda operar en el vacío. Para no convertirse en una técnica algorítmica de gestión del riesgo, debe integrarse en una arquitectura institucional de justicia temprana que articule la capacidad de detección de la IA con mecanismos humanos de información, revisión y solución de controversias, sin que la tecnología cierre por sí misma el conflicto ni sustituya los espacios deliberativos propios del derecho.

En esa arquitectura, los mecanismos adecuados de solución de controversias adquieren una función renovada, porque la mediación, la negociación asistida o la conciliación no deberían concebirse solo como vías alternativas una vez surgida la disputa, sino también como instrumentos de intervención temprana cuando aparecen señales de deterioro relacional o incumplimientos repetidos. Se establece así un reparto en el que la IA contribuye a detectar y orientar, mientras que estos mecanismos aportan escucha, reconocimiento y construcción cooperativa de soluciones (Menkel-Meadow 1995, 2665–2670, Barona Vilar 2018, 27, Soleto Muñoz 2021, Gonzalo Quiroga 2024, 152–156).

El legislador español ha dado un paso en esta dirección con la Ley Orgánica 1/2025, que configura como requisito de procedibilidad en el orden civil el intento previo de acudir a un medio adecuado de solución de controversias (Ley Orgánica 1/2025, art. 5). Esa reforma adelanta el momento en que el sistema ofrece cauces consensuales, aunque sigue activándose cuando el conflicto ya existe, y es ahí donde la IA podría contribuir a adelantarlo todavía más, identificando las situaciones en las que conviene ofrecer esos cauces antes de que el desacuerdo cristalice.

La clave está en diseñar itinerarios progresivos, que comiencen con información clara sobre derechos y opciones disponibles, continúen, cuando sea necesario, con revisión humana, negociación o mediación, y mantengan abierta la tutela judicial ante una especial complejidad o una afectación intensa de derechos. El derecho anticipativo no sustituye, así, la arquitectura tradicional de justicia, sino que la reordena en el tiempo para intervenir antes y con mayor capacidad de evitar la escalada del conflicto. Esta integración resulta esencial para impedir que la prevención inteligente se utilice para filtrar reclamaciones o presionar hacia acuerdos rápidos, y que la tecnología privatice de forma opaca la gestión del conflicto. Frente a ambos riesgos, el derecho anticipativo debe reforzar la posición de las partes y facilitar soluciones tempranas sin renunciar a las garantías de igualdad, revisión humana y acceso a la tutela judicial efectiva.

Una arquitectura de este tipo requiere, además, identificar a sus actores. En el ámbito privado, los primeros llamados a intervenir serían las plataformas y grandes empresas que

ya disponen de datos sobre reclamaciones, siempre que la detección se acompañe de cauces independientes de revisión. En el ámbito público, por su parte, las administraciones con alto volumen de recursos, los servicios de orientación jurídica y las oficinas de consumo podrían emplear la IA para identificar patrones de conflictividad y ofrecer información o mediación tempranas.

A las autoridades de supervisión les correspondería verificar el cumplimiento de las garantías, y a los profesionales de la mediación y la abogacía, intervenir en la fase de respuesta para que la alerta se traduzca en diálogo y no en una decisión unilateral. Ninguno de estos actores debería operar, en todo caso, sin una evaluación previa de impacto en los derechos, sin trazabilidad de las alertas y de sus resultados y sin la posibilidad real de que la persona afectada acuda a una instancia independiente.

5. Riesgos y límites del derecho anticipativo

Definidos los principios y el marco normativo del derecho anticipativo, conviene detenerse en aquello que puede torcer su propósito, puesto que, al desplazar la intervención jurídica hacia fases previas del conflicto, su potencial innovador coincide con su principal amenaza, y intervenir antes puede significar proteger antes, pero también controlar antes. Zuboff ha mostrado cómo la economía de los datos convierte la experiencia humana en materia prima para predecir y modificar conductas (Zuboff 2019, 196–203), y Brownsword advierte del desplazamiento de la regulación normativa por una gestión tecnológica del entorno que reduce el espacio de elección (Brownsword 2019, 19–28). Si el derecho anticipativo se utilizara para vigilar, perfilar o neutralizar reclamaciones, dejaría de ser una teoría de prevención jurídica para convertirse en una técnica de administración preventiva del riesgo.

Por eso, los subapartados siguientes examinan esos riesgos y, por último, los límites que derivan de la propia naturaleza de la predicción algorítmica.

5.1. Sesgo y anticipación discriminatoria

El primero de esos riesgos aparece en la propia lógica predictiva, porque los sistemas de IA trabajan sobre datos pasados, lo que les permite identificar regularidades, pero también reproducir sesgos institucionales o prácticas discriminatorias ya incorporadas a esos datos. Barocas y Selbst han mostrado que la minería de datos puede producir efectos discriminatorios incluso sin intención de discriminar, al heredar los prejuicios presentes en los datos de entrenamiento (Barocas y Selbst 2016, 671–689), y O'Neil y Eubanks han documentado cómo los modelos opacos aplicados a gran escala castigan con especial intensidad a los colectivos más vulnerables (O'Neil 2016, 3–15, Eubanks 2018).

Este riesgo resulta especialmente grave en la prevención de conflictos, porque la intervención se produce antes de que exista una controversia formalizada, y si un sistema identifica como «riesgo» a determinados perfiles de consumidores, usuarios o solicitantes, puede activar medidas desfavorables sin que la persona haya incurrido en conducta reprochable alguna, transformando la prevención en anticipación discriminatoria. El

problema tiene, además, una dimensión jurídica más profunda, porque obliga a preguntarse qué se considera riesgo y qué consecuencias se asocian a una clasificación. Wachter ha advertido, en este sentido, que los grupos creados por los algoritmos no suelen coincidir con las categorías protegidas por el derecho antidiscriminatorio, lo que deja a sus miembros sin tutela efectiva (Wachter 2022). En un modelo anticipativo, en suma, el sesgo opera antes del daño visible y antes de la contradicción, y sus efectos resultan por ello más difíciles de detectar e impugnar (Selbst *et al.* 2019, 59–65).

5.2. Vigilancia y perfilado

A este riesgo se añade la posibilidad de que la prevención inteligente derive en vigilancia, porque la acumulación de datos para identificar señales tempranas de conflictividad puede convertirse en monitorización constante, y el problema no reside solo en la cantidad de datos tratados, sino en la finalidad para la que se utilizan (Troncoso Reigada 2021, Cerrillo i Martínez 2021, 1–9). Si la prevención del conflicto se apoyara en una vigilancia generalizada, el nuevo paradigma nacería debilitado desde el punto de vista del Estado de derecho. Por eso la recogida de datos con fines preventivos solo es legítima cuando se vincula a finalidades concretas, límites temporales y posibilidades reales de revisión, como exige el principio de limitación de la finalidad.

Estrechamente unido a la vigilancia está el perfilado. En un modelo anticipativo mal diseñado, la intervención puede apoyarse en probabilidades o semejanzas estadísticas con otros casos y generar una forma de responsabilidad anticipada, en la que no se actúa porque la persona haya incumplido, sino porque el sistema estima que podría hacerlo. En las plataformas digitales el perfilado puede afectar al acceso al servicio o a la reputación, en el ámbito laboral a la evaluación de trabajadores y en la Administración a inspecciones o controles, y en todos los casos la anticipación puede reforzar la eficiencia, pero también trasladar a la persona afectada la carga de demostrar que no constituye un riesgo.

Por eso Citron y Pasquale han reclamado garantías de debido proceso frente a las puntuaciones y predicciones automatizadas (Citron y Pasquale 2014, 8–18), Zarsky ha mostrado que la transparencia de los modelos predictivos es necesaria, pero no basta para neutralizar todos sus riesgos (Zarsky 2013, 1508–1516), y Crawford sitúa estas clasificaciones en una economía política más amplia de extracción de datos y concentración de poder (Crawford 2021, 122–130).

5.3. Brecha digital y exclusión

La brecha digital constituye un límite de naturaleza distinta, pero igualmente relevante, porque un modelo preventivo basado en interfaces digitales o en clasificación algorítmica puede beneficiar a quienes disponen de recursos y habilidades suficientes y desplazar a quienes ya se encuentran en posiciones de vulnerabilidad (Piñar Mañas 2018, 95–112, Caro Catalán 2022, 110–124, Sánchez Martínez 2025, 5–12). Se trata de una brecha que no es solo tecnológica, sino también lingüística y cognitiva, y que entraña el riesgo de que el derecho anticipativo mejore la gestión del conflicto para quienes ya acceden al sistema mientras invisibiliza a quienes más necesitan una intervención temprana. Por eso el acceso temprano a la justicia debe diseñarse desde una lógica inclusiva, con canales

alternativos, asistencia humana y lenguaje claro (Segura 2023, 45–72, Polo Olivarez 2025, 381–393).

5.4. Los límites de la anticipación algorítmica: fiabilidad, falsos positivos y sobreconfianza

A estos riesgos, derivados de un uso indebido de la anticipación, se suman límites inherentes a la propia predicción. El más evidente es que toda predicción algorítmica es probabilística y, por tanto, yerra en dos direcciones, puesto que los falsos positivos pueden activar intervenciones innecesarias o estigmatizantes y los falsos negativos dejan sin detectar conflictos reales y generan una falsa apariencia de control. Cuando el fenómeno que se quiere predecir es poco frecuente, incluso un sistema con una tasa de acierto elevada puede producir más alertas falsas que verdaderas. El volumen de intervenciones injustificadas no depende, por tanto, solo de la calidad técnica del modelo, sino del umbral que se fije para actuar, y ese umbral, lejos de ser una decisión técnica, es una decisión normativa que debe estar gobernada por la proporcionalidad.

A ello se añade que los sistemas predicen a partir de correlaciones observadas en el pasado sin comprender las razones del conflicto (Mittelstadt *et al.* 2016, 3–5), de forma que, cuando cambian las circunstancias, el comportamiento de los usuarios o las propias reglas, el modelo puede perder fiabilidad sin que nadie lo advierta. La intervención anticipativa altera, además, el fenómeno que mide, porque si el sistema evita un conflicto nunca se sabrá con certeza si este se habría producido, lo que dificulta evaluar su eficacia y facilita atribuirle éxitos que no le corresponden.

Existe, por último, el riesgo de atribuir a la tecnología una capacidad de anticipación superior a la que realmente posee. La supervisión humana puede degradarse en mera ratificación cuando quien revisa confía en exceso en la recomendación automatizada, y precisamente por ello el Reglamento de IA exige que las personas encargadas de la supervisión sean conscientes de ese sesgo de automatización (Parlamento Europeo y Consejo 2024a, art. 14.4.b). En esa misma línea, Kroll y otros autores han propuesto herramientas técnicas y procedimentales que permitan verificar que los sistemas automatizados se comportan como se declara (Kroll *et al.* 2017). Para el derecho anticipativo, la consecuencia es que la alerta algorítmica debe tratarse como un indicio que justifica, como máximo, una actuación de información o de revisión, y nunca como prueba de un conflicto ni como fundamento suficiente de una restricción de derechos.

5.5. Compatibilidad con el Estado de derecho

Todos estos riesgos convergen en la cuestión de la compatibilidad del derecho anticipativo con el Estado de derecho, puesto que la anticipación jurídica no puede construirse al margen de los principios de legalidad, igualdad, proporcionalidad y tutela judicial efectiva. Boix Palop ha sostenido que los algoritmos que la Administración emplea para decidir cumplen una función equivalente a la de los reglamentos y deben someterse a garantías análogas (Boix Palop 2020, 245–258), y De Gregorio, desde el constitucionalismo digital, reclama extender los límites constitucionales al poder privado de las plataformas (De

Gregorio 2022). Ambas perspectivas apuntan en la misma dirección, porque si la IA se utiliza para adoptar decisiones preventivas opacas o basadas en criterios no verificables, el modelo anticipativo puede erosionar las mismas garantías que pretende fortalecer. De ahí la importancia de incorporar esas garantías desde el diseño de los sistemas y no como corrección posterior (Cotino Hueso 2019, 29–48).

El límite esencial del derecho anticipativo es, en suma, que nunca debe convertir la prevención en sustituto de la garantía, porque la detección temprana no puede reemplazar la prueba ni la eficiencia justificar la opacidad. Su finalidad debe ser justamente la contraria, la de hacer visibles los conflictos emergentes, reforzar la posición de las partes y abrir itinerarios de solución sin abandonar las garantías que justifican su legitimidad, y por eso su éxito no debería medirse únicamente por la reducción de litigios, sino por la calidad de las respuestas tempranas y la mejora real del acceso a la justicia.

6. Conclusiones

Incorporar la IA al ámbito jurídico obliga a plantear una cuestión que va más allá de la eficiencia tecnológica o de la asistencia a los operadores jurídicos, porque la pregunta decisiva no es solo si la IA puede ayudar a resolver más rápido los conflictos, sino si puede modificar el momento en que el derecho interviene frente a ellos. En torno a esa pregunta se ha construido la tesis central de este trabajo, según la cual la mayor aportación de la IA al derecho no consiste en automatizar la decisión jurídica, sino en transformar la temporalidad de la intervención jurídica frente al conflicto.

El modelo jurídico tradicional se ha construido sobre una lógica reactiva que ha garantizado seguridad jurídica y tutela judicial efectiva, pero que muestra límites relevantes en sociedades digitales marcadas por la rapidez, la masividad y la asimetría informativa, y es en ese contexto donde la IA abre la posibilidad de identificar señales tempranas de conflictividad antes de que el conflicto se consolide como litigio.

A partir de esta constatación, el trabajo ha propuesto la categoría de derecho anticipativo, no como una nueva rama del ordenamiento, sino como una función jurídica emergente que consiste en intervenir en fases previas del conflicto mediante mecanismos de detección, diagnóstico y respuesta temprana. Lejos de sustituir al derecho reactivo o de desplazar la tutela judicial, lo complementa con una capa previa de prevención inteligente que permite al derecho llegar antes y activar respuestas proporcionadas antes de que el conflicto escale.

La anticipación no puede convertirse, sin embargo, en una nueva forma de control, y por eso este trabajo ha defendido que el derecho anticipativo solo es jurídicamente admisible si se articula sobre principios estrictos (anticipación, proporcionalidad, supervisión humana, transparencia, confianza y acceso temprano a la justicia) que permitan diferenciar la prevención legítima de una mera gestión algorítmica del riesgo. Tampoco debería medirse únicamente por su capacidad de reducir litigios, porque una disminución de controversias no equivale necesariamente a mayor justicia si también significa menor visibilidad del conflicto o más dificultad para impugnar actuaciones opacas. El criterio de éxito debe ser, en consecuencia, cualitativo y atender a la calidad de la información, a la prontitud de la intervención y a la preservación efectiva de los derechos.

De ello se desprenden las condiciones en que la prevención inteligente puede convertirse en una práctica jurídicamente legítima y socialmente deseable. Necesita, ante todo, una base normativa y una finalidad definida, que vincule la detección a la protección de derechos y no solo al interés de quien controla el sistema. Esa finalidad debe traducirse en umbrales de intervención proporcionados, de manera que las alertas activen preferentemente medidas de información, revisión o derivación, y nunca restricciones de derechos sin una decisión humana motivada. Requiere también transparencia frente a la persona afectada y trazabilidad de las alertas, de sus resultados y de sus errores, incluida la tasa de falsos positivos, algo que solo puede verificarse mediante evaluaciones de impacto previas y auditorías periódicas sobre sesgos y efectos en colectivos vulnerables, acompañadas de canales alternativos no digitales. Y exige, en todo caso, preservar la vía judicial y los mecanismos independientes de revisión. El marco europeo vigente ofrece ya anclajes para cada una de estas condiciones, aunque de forma dispersa.

El desplazamiento temporal transforma también la función del derecho, cuya dimensión preventiva deja de limitarse a establecer normas generales y deberes de diligencia para incluir una intervención individualizada sobre conflictos que todavía no han sido reclamados. Esta transformación genera una tensión con la autonomía individual, porque el derecho tradicional deja en manos de la persona la decisión de nombrar, atribuir y reclamar un agravio, mientras que el derecho anticipativo puede adelantarse a esa decisión. Por eso la respuesta anticipativa debe ser, por regla general, una oferta de información, de revisión o de diálogo, y no una imposición, y debe respetar también la libertad de no reclamar y de no ser objeto de una atención que no se ha solicitado.

En este sentido, la IA puede funcionar como infraestructura de gobernanza preventiva y permitir que instituciones y plataformas aprendan de los conflictos que generan y corrijan los diseños que los producen, siempre que esa infraestructura permanezca jurídicamente gobernada, porque la tecnología puede identificar patrones, pero corresponde al derecho decidir qué riesgos son relevantes y qué límites no pueden superarse.

Conviene advertir, por último, que este trabajo tiene un alcance fundamentalmente conceptual y que los ejemplos empleados ilustran la tesis sin permitir medir los efectos reales de la anticipación. Quedan abiertas, por ello, varias líneas de investigación empírica, como el estudio de la fiabilidad y de las tasas de error de los sistemas de detección temprana en contextos jurídicos concretos, la evaluación de experiencias piloto de orientación y mediación tempranas en servicios públicos de consumo o de justicia, y el análisis de cómo perciben las personas afectadas una intervención jurídica que se adelanta a su propia reclamación.

El derecho anticipativo no propone un derecho sin conflictos, sino un derecho capaz de reconocer antes las condiciones que los generan y de abrir vías tempranas de solución sin renunciar a las garantías del Estado de derecho. Su aportación consiste en desplazar la pregunta jurídica desde el “cómo resolver el conflicto una vez producido” hacia el “cómo evitar que determinadas tensiones lleguen innecesariamente a convertirse en litigios”. En ese desplazamiento temporal, más que en cualquier promesa de automatización, se encuentra la verdadera innovación jurídica de la IA, y también la medida en que este trabajo espera haber contribuido a delimitarla con el rigor y las cautelas que exige el Estado de derecho.

Referencias

Airbnb, 2026. *How Airbnb screens for potential disruptive gathering risk*. Disponible en: https://www.airbnb.co.uk/help/article/3280

Amazon, 2024. *How Amazon is using AI to spot fake reviews*. Disponible en: https://www.aboutamazon.com/news/policy-news-views/how-ai-spots-fake-reviews-amazon

Barocas, S., y Selbst, A. D., 2016. Big data's disparate impact. *California Law Review*, *104*(3), 671–732.

Barona Vilar, S., 2018. *Nociones y principios de las ADR*. Valencia: Tirant lo Blanch.

Battaglini, M., y Rasmussen, S., 2019. Transparency, automated decision-making processes and personal profiling. *Journal of Data Protection & Privacy*, *2*(4), 331–349. Disponible en: https://doi.org/10.69554/STNB6990

Blackaby, N., *et al.*, 2022. *Redfern and Hunter on international arbitration*. 7ª ed. Oxford University Press.

Bloch-Wehba, H., 2020. Access to algorithms. *Fordham Law Review*, *88*(4), 1265–1314. Disponible en: https://ir.lawnet.fordham.edu/flr/vol88/iss4/2

Boix Palop, A., 2020. Los algoritmos son reglamentos: la necesidad de extender las garantías propias de las normas reglamentarias a los programas empleados por la Administración para la adopción de decisiones. *Revista de Derecho Público: Teoría y Método*, 1, 223–269. https://doi.org/10.37417/RPD/vol_1_2020_33

Bolton, R. J., y Hand, D. J., 2002. Statistical fraud detection: A review. *Statistical Science*, *17*(3), 235–255. https://doi.org/10.1214/ss/1042727940

Bourdieu, P., 1987. The force of law: Toward a sociology of the juridical field. *Hastings Law Journal*, *38*(5), 805–853. Disponible en: https://repository.uclawsf.edu/hastings_law_journal/vol38/iss5/3/

Brownsword, R., 2019. *Law, technology and society: Re-imagining the regulatory environment*. Abingdon: Routledge.

Burrell, J., 2016. How the machine “thinks”: Understanding opacity in machine learning algorithms. *Big Data & Society*, *3*(1), 1–12. https://doi.org/10.1177/2053951715622512

Calo, R., y Rosenblat, A., 2017. The taking economy: Uber, information, and power. *Columbia Law Review*, *117*(6), 1623–1690.

Cappelletti, M., y Garth, B., 1978. *Access to justice: The worldwide movement to make rights effective. A general report*. Alphen aan den Rijn: Sijthoff and Noordhoff.

Caro Catalán, J., 2022. Brechas digitales y acceso a la Justicia: el caso de España. *En*: M. M. Bustamante Rúa, ed., *Justicia y Sociedad 5.0*. Medellín: Fondo Editorial IUE, 110–124.

Castets-Renard, C., 2019. Accountability of algorithms in the GDPR and beyond: A European legal framework on automated decision-making. *Fordham Intellectual Property, Media & Entertainment Law Journal*, *30*(1), 91–137. Disponible en: https://ir.lawnet.fordham.edu/iplj/vol30/iss1/3

Cerrillo i Martínez, A., 2021. Retos y oportunidades del uso de la inteligencia artificial en las administraciones públicas. *Revista Iberoamericana de Gobierno Local*, 19, 1–9. Disponible en: https://rigl.s3.amazonaws.com/media/documents/19/agusti-cerrillo-i-martinez_retos-y-oportunidades-del-uso-de-la-inteligencia-artificial-en-las-administraciones-publicas_frSXByQ.pdf

Chauhan, A., 2020. Towards the systemic review of automated decision-making systems. *Judicial Review*, *25*(4), 285–295. Disponible en: https://doi.org/10.1080/10854681.2020.1871714

Citron, D. K., y Pasquale, F., 2014. The scored society: Due process for automated predictions. *Washington Law Review*, *89*(1), 1–33.

Coglianese, C., y Lehr, D., 2017. Regulating by robot: Administrative decision making in the machine-learning era. *Georgetown Law Journal*, 105, 1147–1223.

Corbett, J. R., 2016. At the nexus of big data and dispute resolution: A case study on conflict-related search data. *Mediation Theory and Practice*, *1*(1), 5–33.

Cotino Hueso, L., 2019. Ética en el diseño para el desarrollo de una inteligencia artificial, robótica y big data confiables y su utilidad desde el derecho. *Revista Catalana de Dret Públic*, 58, 29–48. Disponible en: https://doi.org/10.2436/rcdp.i58.2019.3303

Crawford, K., 2021. *Atlas of AI: Power, politics, and the planetary costs of artificial intelligence*. New Haven: Yale University Press.

Das, D., 2025. *Fortify your business using PayPal's risk and fraud management solutions*. PayPal. Disponible en: https://developer.paypal.com/community/blog/paypal-fraud-risk-management-solutions/

De Gregorio, G., 2022. *Digital constitutionalism in Europe: Reframing rights and powers in the algorithmic society*. Cambridge University Press.

Dellarocas, C., 2003. The digitization of word of mouth: Promise and challenges of online feedback mechanisms. *Management Science*, *49*(10), 1407–1424.

Ester Sánchez, A. T., 2025. La inteligencia artificial en la justicia. Desafíos y oportunidades en la toma de decisiones judiciales. *Anales de la Cátedra Francisco Suárez*, 59, 317–340. Disponible en: https://doi.org/10.30827/acfs.v59i.31404

Eubanks, V., 2018. *Automating inequality: How high-tech tools profile, police, and punish the poor*. Nueva York: St. Martin's Press.

Felstiner, W. L. F., Abel, R. L., y Sarat, A., 1981. The emergence and transformation of disputes: Naming, blaming, claiming. *Law & Society Review*, *15*(3–4), 631–654. Disponible en: https://doi.org/10.2307/3053505

Fuller, L. L., 1978. The forms and limits of adjudication. *Harvard Law Review*, *92*(2), 353–409.

Gillespie, T., 2018. *Custodians of the Internet: Platforms, content moderation, and the hidden decisions that shape social media*. New Haven: Yale University Press.

González Fernández, A. I., 2022. La irrupción de la inteligencia artificial en la resolución alternativa de conflictos. *En:* S. Calaza López y M. Llorente Sánchez-Arjona, eds., *Inteligencia artificial legal y administración de justicia*. Cizur Menor: Thomson Reuters Aranzadi, 543-553.

González Fernández, A. I., 2024. Retos y oportunidades de la inteligencia artificial en el campo de la resolución de conflictos. *Revista de Derecho UCLAEH*, *3*(3), 51–64. Disponible en: https://doi.org/10.70640/rdclaeh.3.3.4

Gonzalo Quiroga, M., 2024. Estudio jurídico de la inteligencia artificial en la justicia alternativa: educación, ODS y propuesta humanista. *Revista de Educación y Derecho* [en línea], número extraordinario 2, 147–182. Disponible en: https://doi.org/10.1344/REYD2024.2-Extraordinario.49176

Google, 2026a. *Requests to delist content under European privacy law*. Disponible en: https://transparencyreport.google.com/eu-privacy/overview

Google, 2026b. *European privacy requests: Search removals FAQs*. Disponible en: https://support.google.com/transparencyreport/answer/7347822

Gorwa, R., Binns, R., y Katzenbach, C., 2020. Algorithmic content moderation: Technical and political challenges in the automation of platform governance. *Big Data & Society* [en línea], *7*(1), 1–15. Disponible en: https://doi.org/10.1177/2053951719897945

High-Level Expert Group on Artificial Intelligence, 2019. *Ethics guidelines for trustworthy AI*. Brussels: European Commission.

Hodges, C., 2019. *Delivering dispute resolution: A holistic review of models in England and Wales*. Oxford: Hart.

Kaminski, M. E., 2020. Understanding transparency in algorithmic accountability. *En:* W. Barfield, ed., *The Cambridge handbook of the law of algorithms*. Cambridge University Press, 121–138.

Katsh, E., Schwartz-Maor, T., y Rabinovich-Einy, O., 2025. AI and the future of disputing: Naming, blaming, claiming, and preventing. *Conflict Resolution Quarterly*, *43*(1), 79–89. Disponible en: https://doi.org/10.1002/crq.21477

Katsh, E., y Rabinovich-Einy, O., 2017. *Digital justice: Technology and the Internet of disputes*. Oxford University Press.

Kroll, J. A., *et al*., 2017. Accountable algorithms. *University of Pennsylvania Law Review*, *165*(3), 1-66.

Lagioia, F., *et al*., 2022. AI in search of unfairness in consumer contracts: The terms of service landscape. *Journal of Consumer Policy*, *45*(3), 481–536. Disponible en: https://doi.org/10.1007/s10603-022-09520-9

Larsson, S., 2017. All-seeing giants and blindfolded dwarfs: On information-asymmetries on data-driven markets. *En:* European Liberal Forum, *New economic models: Tools for political decision makers dealing with the changing European economies*. Brussels: European Liberal Forum, 66-78. Disponible en: https://portal.research.lu.se/en/publications/all-seeing-giants-and-blindfolded-dwarfs-on-information-asymmetri

Luca, M., y Zervas, G., 2016. Fake it till you make it: Reputation, competition, and Yelp review fraud. *Management Science*, *62*(12), 3412–3427.

Luhmann, N., 2004. *Law as a social system*. Trans. K. Ziegert. Oxford University Press.

Marcos Francisco, D., 2023. Smart ODR y su puesta en práctica: el salto a la inteligencia artificial. *Revista General de Derecho Procesal*, 59, 1–41.

Mayzlin, D., Dover, Y., y Chevalier, J., 2014. Promotional reviews: An empirical investigation of online review manipulation. *American Economic Review*, *104*(8), 2421–2455.

Medina Morales, D., 2023. La labor del jurista en la resolución del conflicto social. El Derecho como sistema de comunicación y de entendimiento social. *Millars. Espai i Història*, *54*(1), 95–116. Disponible en: https://doi.org/10.6035/Millars.2023.54.5

Mehta, D., 2025. *How Amazon uses AI innovations to stop fraud and counterfeits*. Amazon. Disponible en: https://www.aboutamazon.com/news/policy-new-views/amazon-brand-protection-report-2024-counterfeit-products

Menkel-Meadow, C., 1991. Pursuing settlement in an adversary culture: A tale of innovation co-opted or “The Law of ADR”. *Florida State University Law Review*, *19*(1), 1–46. Disponible en: https://ir.law.fsu.edu/cgi/viewcontent.cgi?article=1668&context=lr

Menkel-Meadow, C., 1995. Whose dispute is it anyway?: A philosophical and democratic defense of settlement (in some cases). *Georgetown Law Journal*, 83, 2663–2696.

Meta, 2023. *How technology detects violations*. Disponible en: https://transparency.meta.com/enforcement/detecting-violations/technology-detects-violations/

Micklitz, H. W., 2021. Digital architecture of private law relations. *En:* S. Grundmann, H. W. Micklitz y M. Renner, *New private law theory: A pluralist approach*. Cambridge University Press, 298–314. Disponible en: https://doi.org/10.1017/9781108760089.018

Miller, R. E., y Sarat, A., 1981. Grievances, claims, and disputes: Assessing the adversary culture. *Law & Society Review*, *15*(3–4), 525–566. Disponible en: https://doi.org/10.2307/3053502

Mittelstadt, B. D., *et al*., 2016. The ethics of algorithms: Mapping the debate. *Big Data & Society*, *3*(2), 1–21. Disponible en: https://doi.org/10.1177/2053951716679679

Moore, C. W., 2014. *The mediation process: Practical strategies for resolving conflict*. 4ª ed. San Francisco: Jossey-Bass.

Moreno Buján, M., 2021. Evolución antropojurídica y sociojurídica de la mediación y las prácticas reconciliatorias: una mirada desde distintas culturas y costumbres a partir de la perspectiva del pluralismo jurídico. *Anuario Iberoamericano de Derecho Internacional Penal*, 8, 1–36. Disponible en: https://doi.org/10.12804/revistas.urosario.edu.co/anidip/a.10225

Ngai, E. W. T., *et al*., 2011. The application of data mining techniques in financial fraud detection: A classification framework and an academic review of literature. *Decision Support Systems*, *50*(3), 559–569.

O'Neil, C., 2016. *Weapons of math destruction: How big data increases inequality and threatens democracy*. Nueva York: Crown.

Ortiz Laverde, S. M., y Ortiz Cabrera, M. A., 2025. Soberanía asimétrica: El derecho administrativo ante el desafío de regular el espacio digital. *Ius et Veritas*, (71), 1–16. Disponible en: https://doi.org/10.18800/iusetveritas.202502.003

Pasquale, F., 2015. *The black box society: The secret algorithms that control money and information*. Cambridge, MA: Harvard University Press.

PayPal, 2024. *Harnessing machine learning fraud detection technologies*. Disponible en: https://www.paypal.com/us/brc/article/payment-fraud-detection-machine-learning

Piñar Mañas, J. L., 2018. Identidad y persona en la sociedad digital. *En:* T. de la Quadra-Salcedo y J. L. Piñar Mañas, eds., *Sociedad Digital y Derecho*. Madrid: BOE, 95–112.

Polo Olivarez, I. A., 2025. Brechas digitales e inteligencia artificial: Retos en el acceso a la justicia. *Forseti. Revista de Derecho*, *14*(22), 381–393. Disponible en: https://doi.org/10.21678/forseti.v14i22.2842

Rabinovich-Einy, O., y Katsh, E., 2012. Technology and the future of dispute systems design. *Harvard Negotiation Law Review*, 17, 151–199. Disponible en: https://journals.law.harvard.edu/hnlr/wp-content/uploads/sites/91/2012/09/151-200.pdf

Rosenblat, A., y Stark, L., 2016. Algorithmic labor and information asymmetries: A case study of Uber's drivers. *International Journal of Communication*, 10, 3758–3784.

Sánchez Martínez, M. O., 2025. Brechas digitales y sesgos en la inteligencia artificial. ¿Una nueva forma de discriminación? *Universitas. Revista de Filosofía, Derecho y Política*, 47, 5–38. Disponible en: https://doi.org/10.20318/universitas.2025.9573

Sander, F. E. A., 1976. Varieties of dispute processing. *Federal Rules Decisions*, 70, 321–326.

Schoinas, T., y Qadir, G., 2017. A feasibility experiment on the application of predictive coding to instant messaging corpora. *2017 IEEE International Conference on Big Data (Big Data)*, 11–14. December, Boston, 3835–3840. Disponible en: https://doi.org/10.1109/BigData.2017.8258386

Schreuer, C., 2008. What is a legal dispute? *En:* I. Buffard *et al*., eds., *International law between universalism and fragmentation: Festschrift in honour of Gerhard Hafner*. Leiden: Brill, 959–980.

Segura, R., 2023. Inteligencia artificial y administración de justicia: desafíos derivados del contexto latinoamericano. *Revista de Bioética y Derecho*, (58), 45–72. Disponible en: https://doi.org/10.1344/rbd2023.58.40601

Selbst, A. D., *et al*., 2019. Fairness and abstraction in sociotechnical systems. *Proceedings of the Conference on Fairness, Accountability, and Transparency*. Nueva York: ACM, 59–68.

Solar Cayón, J. I., 2020. *La inteligencia artificial jurídica: nuevas herramientas y perspectivas metodológicas para el jurista*. Cizur Menor: Thomson Reuters Aranzadi.

Soleto Muñoz, H., 2021. *Mediación y resolución de conflictos: técnicas y ámbitos*. Madrid: Tecnos.

Surden, H., 2014. Machine learning and law. *Washington Law Review*, *89*(1), 87–115.

Susskind, R., 2019. *Online courts and the future of justice*. Oxford University Press.
Troncoso Reigada, A., ed., 2021. *Comentario al Reglamento General de Protección de Datos y a la Ley Orgánica de Protección de Datos personales y garantía de los derechos digitales*. Cizur Menor: Thomson Reuters Aranzadi.
Uber, 2026. *Understanding why drivers and delivery people can lose access to their accounts*. Disponible en: https://www.uber.com/us/en/drive/driver-app/deactivation-review/
Unesco, 2021. *Recommendation on the ethics of artificial intelligence*. París: Unesco.
Vilalta Nicuesa, A. E., y Gili Saldaña, M., 2025. AI-driven alternative and online dispute resolution in the European Union: An analysis of the legal framework and a proposed categorization. *Computer Law & Security Review*, 57, article 106145. Disponible en: https://doi.org/10.1016/j.clsr.2025.106145
Wachter, S., 2022. The theory of artificial immutability: Protecting algorithmic groups under anti-discrimination law. *Tulane Law Review*, *97*(2), 1-48.
Wachter, S., Mittelstadt, B., y Floridi, L., 2017. Why a right to explanation of automated decision-making does not exist in the General Data Protection Regulation. *International Data Privacy Law*, *7*(2), 76–99.
Yeung, K., 2018. Algorithmic regulation: A critical interrogation. *Regulation & Governance*, *12*(4), 505–523.
YouTube, 2026. *Community Guidelines enforcement*. Disponible en: https://transparencyreport.google.com/youtube-policy/overview
Zarsky, T. Z., 2013. Transparent predictions. *University of Illinois Law Review*, *2013*(4), 1503–1569.
Zuboff, S., 2019. *The age of surveillance capitalism: The fight for a human future at the new frontier of power*. Nueva York: PublicAffairs.

Fuentes normativas y jurisprudenciales

Ley 2/2023, de 20 de febrero, reguladora de la protección de las personas que informen sobre infracciones normativas y de lucha contra la corrupción. *Boletín Oficial del Estado*, núm. 44, 21 de febrero de 2023.
Ley 31/1995, de 8 de noviembre, de Prevención de Riesgos Laborales. *Boletín Oficial del Estado*, núm. 269, 10 de noviembre de 1995.
Ley 7/2022, de 8 de abril, de residuos y suelos contaminados para una economía circular. *Boletín Oficial del Estado*, núm. 85, 9 de abril de 2022.
Ley Orgánica 1/2025, de 2 de enero, de medidas en materia de eficiencia del Servicio Público de Justicia. *Boletín Oficial del Estado*, núm. 3, 3 de enero de 2025.
Parlamento Europeo y Consejo, 2016. Reglamento (UE) 2016/679, 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 (Reglamento General de Protección de Datos). *Diario Oficial de la Unión Europea*, L 119, 4 de mayo de 2016.
Parlamento Europeo y Consejo, 2022. Reglamento (UE) 2022/2065, de 19 de octubre de 2022, relativo a un mercado único de servicios digitales (Reglamento de Servicios Digitales). *Diario Oficial de la Unión Europea*, L 277, 27 de octubre de 2022.
Parlamento Europeo y Consejo, 2024a. Reglamento (UE) 2024/1689, de 13 de junio de 2024, por el que se establecen normas armonizadas en materia de inteligencia artificial. *Diario Oficial de la Unión Europea*, L, 12 de julio de 2024.
Parlamento Europeo y Consejo, 2024b. Directiva (UE) 2024/2831, de 23 de octubre de 2024, relativa a la mejora de las condiciones laborales en el trabajo en plataformas. *Diario Oficial de la Unión Europea*, L, 11 de noviembre de 2024.
Real Decreto Legislativo 1/2007, de 16 de noviembre, por el que se aprueba el texto refundido de la Ley General para la Defensa de los Consumidores y Usuarios y otras leyes complementarias. *Boletín Oficial del Estado*, núm. 287, 30 de noviembre de 2007.
Tribunal de Justicia de la Unión Europea (TJUE), 2014. Sentencia de 13 de mayo de 2014, *Google Spain SL y Google Inc. contra Agencia Española de Protección de Datos y Mario Costeja González*, asunto C-131/12, ECLI:EU:C:2014:317.