REVISTA DE **DERECHO PRIVADO:** DEBATES Y TENDENCIAS
Núm. 02 (2025) · ISSN 3087-2332· doi 10.5354/3087-2332.2025.81261
Recibido: 24/10/2025 · Aprobado: 24/11/2025 · Publicado: 31/12/2025

Inteligencia artificial, ¿una "cosa peligrosa"?

Artificial Intelligence, a "Dangerous Thing"?

Alexander Vargas Tinoco
Universidad Externado de Colombia, Colombia

RESUMEN: El presente artículo busca examinar la posibilidad de catalogar la Inteligencia Artificial (IA) como una "cosa peligrosa" dentro del marco conceptual de la responsabilidad civil en Colombia. Se postula que las concepciones tradicionales de "cosa" y "peligrosidad", utilizadas en la doctrina y jurisprudencia colombiana, resultan insuficientes para abordar las complejas particularidades de esta nueva realidad tecnológica. Se analiza cómo la categorización de la IA como "riesgosa" en la reglamentación de la Unión Europea introduce elementos que se distancian de los planteamientos clásicos de la responsabilidad civil. Por consiguiente, se concluye que estos conceptos deben evolucionar en sintonía con los desarrollos prácticos y teóricos de la IA, con el propósito de alcanzar una convergencia regulatoria coherente y necesaria sobre esta materia.

PALABRAS CLAVE: riesgos, inteligencia artificial, cosa peligrosa, peligro, responsabilidad civil.

ABSTRACT: The present paper seeks to examine the feasibility of characterizing Artificial Intelligence (AI) as a "dangerous thing" within the conceptual framework of civil liability in Colombia. It contends that the traditional understandings of "thing" and "dangerousness" developed in Colombian doctrine and jurisprudence are inadequate to address the complex and unprecedented features of this technological phenomenon. The analysis further considers how the European Union's classification of AI as "high-risk" introduces regulatory elements that depart from the classical foundations of civil liability. Accordingly, the paper concludes that these conceptual categories must evolve in tandem with the practical and theoretical developments surrounding AI, with the aim of achieving a coherent and necessary regulatory convergence in this field.

KEYWORDS: risks, artificial intelligence, dangerous thing, periculum, civil liability.

Introducción

El veloz avance de las tecnologías de procesamiento de la información plantea preguntas y retos jurídicos que deben ser enfrentados desde diferentes áreas del Derecho que, además de dar un respuesta sólida a tales desafíos, deben esforzarse por hacerlo de manera oportuna, *so pena* de quedarse rezagadas respecto de una realidad que cambia velozmente.

En ese contexto, la implementación de la inteligencia artificial (en adelante "IA"), cada vez más difundida y aceptada entre las personas, permite formular la pregunta sobre cuál debería ser el tratamiento conceptual que se le dé a ella desde la responsabilidad civil. En particular, en este escrito me ocuparé de determinar si la IA puede ser considerada una "cosa peligrosa", a la luz de los conceptos de "cosa" y de "peligrosidad" que se manejan a nivel doctrinal y jurisprudencial dentro de este ámbito del Derecho, tomando como referencia el ordenamiento jurídico colombiano. Para tal fin, en la primera parte abordaré de manera preliminar el concepto de IA, para luego examinar, en la segunda y tercera parte, respectivamente, los de "cosa" y "peligrosidad". Esto, con el fin de reflexionar en torno a la posible adecuación de aquella categoría a estas. En síntesis, argumentaré que la manera como ellas se plantean no se ajusta del todo a la realidad de esta tecnología, y añadiré que los criterios para determinar su peligrosidad aún deben ser objeto de desarrollo y reflexión. Sobre este punto, mostraré en la cuarta sección cómo el reglamento actual de la Unión Europea sobre la materia implementa criterios de riesgo diferentes a los tradicionales de la responsabilidad civil, lo que incita a reflexionar sobre la vigencia de estos y la necesidad de armonización o actualización con la nueva regulación, cuestión de la que me ocuparé en la última parte de este trabajo, antes de presentar unas conclusiones.

Del concepto de IA

Antes de considerar qué puede entenderse por IA, es necesario tener presente que hasta el 2020 ciertos organismos multilaterales como la Unión Europea admitían que no existía una definición estándar de este

concepto, sino variados intentos de confeccionar una definición que abarcara los diferentes matices de esta nueva realidad. Así lo reconoció el reporte elaborado por dicha organización en ese año, donde se señala con claridad que hay una tendencia a la sobresimplificación de lo que es la IA, mediante propuestas reduccionistas respecto de lo que ella implica (Samoili, 2020: 7).

Sin embargo, la amplitud del fenómeno y los intentos reduccionistas no deberían llevar al fracaso de los esfuerzos por acotar el concepto. De allí que pueda afirmarse que las definiciones contemporáneas de IA apuntan a relacionar ciertas características para identificar que se trata de un fenómeno con el que podemos tener una interacción "inteligente". En ese sentido, resulta interesante la definición que el Centro de Excelencia para la Inteligencia Artificial de los Estados Unidos, que afirma que la IA "…se refiere a las técnicas computacionales que simulan capacidades cognitivas humanas".[1] Ello muestra que, en tanto seres inteligentes que somos, con la capacidad de autopercibirnos como tales, también somos capaces de percatarnos y de reconocer que alguien o algo ajeno a nosotros goza de habilidades que nos permiten afirmar su inteligencia. Esto es, que gracias a la interacción con tales entidades, denotamos por reflejo sus propias capacidades inteligentes.

Pero, ¿cuáles son esas capacidades? En particular, pensamos que un sistema con capacidad de aprendizaje, de procesamiento de información, de autonomía y de decisión es uno con características que consideraríamos "inteligentes" en su operatividad. En ese sentido, Perez-Ugena señala que:

> lo que define a los sistemas de IA es la capacidad de procesar datos e información de una manera que se asemeja a un comportamiento inteligente, y abarca generalmente aspectos de razonamiento, aprendizaje, creatividad o la capacidad de predecir o planear. Es de un conjunto de técnicas, por tanto, que permiten a un sistema informático simular características que son propias de la inteligencia humana. De manera

1 Government of the United States of America, Centers of Excellence (2025): *Introduction to the AI Guide for Government*. Disponible en https://coe.gsa.gov/coe/ai-guide-for-government/introduction/, consultado por última vez el 23 de octubre de 2025.

> que se automatizan ciertas actividades hasta ahora vinculadas exclusivamente con procesos del pensamiento humano, como pueden ser la toma de decisiones (2024: 311).

A nivel normativo, estas características también están presentes en el Reglamento de la Unión Europea 2024/1689 de 13 de junio de 2024, que además tiene en cuenta que estos sistemas son capaces de funcionar con distintos niveles de autonomía a la hora de procesar información y adaptarse a ella, agregando que sus resultados de salida pueden tener incidencia en el entorno físico o virtual. Así, el artículo 3, numeral 1, del Reglamento señala que un sistema de IA es uno

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

En este marco, es posible sintetizar las características definitorias de la IA, al menos, en las siguientes: artificialidad, aprendizaje, razonamiento, inferencialidad, predictibilidad y autonomía en los procesos de inferencialidad, que pueden emular las capacidades cognitivas humanas y arrojar resultados de salida que pueden incidir en el entorno y las decisiones humanas. Es relevante señalar que la detentación de estas capacidades no necesariamente se traduce en una mera imitación o emulación de las posibilidades humanas de razonamiento, sino que, en veces, puede llegar a superarlas en cuanto a la velocidad y cantidad de procesamiento de la información y la producción de datos de salida.

Con todo, debe destacarse que, a pesar del carácter inteligente de estas tecnologías, algunos denuncian que ellas carecen, hasta el momento, de un elemento necesario para que los resultados de su operación sean totalmente adecuados al contexto humano: la consciencia. En efecto, al decir de Labañino y otros, el término "inteligencia artificial" puede resultar engañoso, en la medida en que sus niveles de comprensión y procesamiento de información no tienen en cuenta

el contexto ni las necesidades de quienes interactúan con ellos. Así, afirman que "la mayoría de los sistemas actuales de IA carecen de una comprensión genuina; operan mediante el reconocimiento de patrones y la manipulación de datos, sin una verdadera percepción del contexto" (Labañino y otros, 2025: 74).

Esta preocupación no necesariamente se resuelve con una programación más compleja o robusta de los sistemas inteligentes, pues, en la orilla contraria, la posibilidad de que ellos sean conscientes del contexto y puedan llegar a ser capaces de percibir su independencia y sus capacidades, también puede resultar preocupante, en particular respecto de lo que puede llegar a implicar esa capacidad respecto de su dependencia de la humanidad. Sobre este punto se volverá más adelante.

En este panorama, la consideración sobre si la IA puede ser comprendida como una "cosa peligrosa" comporta relevancia para determinar cuál es la mejor manera de regular las situaciones en las que ella puede producir daños a otros. El asunto es tan complejo como interesante. Complejo, porque en el uso de la IA intervienen varias partes: desarrolladores de *software*, empresarios productores del hardware o de las cosas que implementan la IA, proveedores de servicios que usan esos productos y los usuarios finales. Ello, aunado a la compleja operación que hay detrás de una IA, que implementa poderosas búsquedas y selecciones de información a través de redes neuronales que se van haciendo mucho más complejas con el aprendizaje que ellas mismas están teniendo gracias a su uso constante o entrenamiento y a la gran extensión y sofisticación que los códigos de programación están adquiriendo para lograr operar de manera autónoma. Además, el tema es interesante porque, en el marco de nuestro Derecho, seguimos navegando a ciegas respecto de esta nueva tecnología, cuyas implementaciones son tan variadas que es sumamente difícil hablar en términos generales.

Con todo, la complejidad de la empresa no puede llevar a la perplejidad del operador jurídico ni del académico, quien tendrá que enfrentar el reto de manera proba. Acá, reitero, solamente se analizará uno de esos retos: el referente a la posible consideración de la IA como una cosa peligrosa. Esta pregunta, por cierto, puede desintegrarse en dos que, metodológicamente, conviene abordar por aparte. La primera, es si la IA puede considerarse una "cosa", y la segunda, si ella es "peligrosa",

esto es, si encaja en los parámetros de peligrosidad que nuestra doctrina y jurisprudencia afirman.

¿"Cosa"?

Inicialmente, la tradición romana conoció el concepto de cosa (*res*) a partir de su corporalidad, esto es, limitado "a los objetos materiales o corpóreos, y no a todos, sino a aquellos que son jurídicamente comerciables" (Iglesias 2010: 154), de manera que cosa era "todo objeto del mundo exterior sobre el cual pueden recaer derechos" (154). Posteriormente, los jurisprudentes efectuaron la distinción entre cosa corporal e incorporal (Gayo Inst. 2, 12-14), que fue recogida por las diferentes codificaciones latinoamericanas, entre ellas, el Código de Bello, del que se derivaron los actuales códigos civiles de Colombia, Chile y Ecuador (art. 653, art. 565 y art. 583 respectivamente).

Así, la idea de "cosa" es amplia, tanto en el lenguaje coloquial como en el jurídico, pues alude, a grandes rasgos, a cualquier entidad corpórea o incorpórea, natural o artificial, real o abstracta, que podamos representarnos mentalemente o percibir con nuestros sentidos (Ternera, 2014: 1). A esta definición hay quien añade que la cosa, en sentido particular, tiene que ver con "todo aquello susceptible de apropiación por el hombre" (Velásquez, 2022: 2), característica que condiciona la posibilidad de que la cosa cambie su naturaleza jurídica a la de "bien", que son aquellas cosas "apropiables" o que pueden ser objeto de apropiación por las personas e integrar su patrimonio personal. En esa línea de pensamiento, se afirma que

> las cosas se consideran bienes, jurídicamente, no solo cuando son útiles al hombre, sino cuando son susceptibles de apropiación: el mar, el aire atmosférico, el sol, son cosas indispensables para la vida terrestre; sin embargo, no son bienes porque no pueden ser objeto de apropiación en provecho de un parti-

> cular, de un ciudadano o de una nación (Planiol y Ripert, 1997: 361).[2]

En esa medida, parte de la doctrina más autorizada afirma que el derecho "solamente se ocupa de las cosas o bienes que pueden ser sometidos al control del hombre y que satisfacen necesidades directas o indirectas" (Ternera, 2014: 1).

Hasta aquí, podría decirse de la IA que se trata de una cosa incorpórea, toda vez que se trata de un sistema codificado resultado de una programación o estrucutra de datos en un lenguaje determinado para procesar información y responder a ella. Por ello, denominaciones alternas a la de "cosa", tales como "artefacto", "máquina" o "mecanismo", no parecen del todo adecuadas para referirse a la IA, pues padecen de un sesgo material, maquinista o fisicista, que no tiene en cuenta su no corporalidad. Así, pues, si bien las IA son formas de *software* que, la mayor de las veces, se implementa en una estructura física (*hardware*) que posibilita la percepción de sus resultados de operación, no sería apropiado confundir su propiedad incorpórea con la materialidad del instrumento del que se vale para operar.

Por otra parte, si se insiste en afirmar que se trata de una "cosa" o un "bien", "a secas", se corre el riesgo de acoger un término que no resulta del todo convincente si se quiere corresponder a la magnitud de las capacidades operativas y de interacción de la IA con la humanidad, o su impacto en nuestro entorno. Esto tiene que ver con el hecho de que, por primera vez en la historia, el ser humano ha creado algo con capacidad inferencial y de razonamiento autónomo, capaz de aprender y de interactuar con él de manera lógica, estructurada y propositiva, con gran potencial de aplicación, fenómeno que ha llevado a afirmar que estamos frente a una quinta revolución industrial, según los efectos que se están evidenciando en la economía y las dinámicas humanas (Akubo, Odiji y Muhammed, 2025).

---

2 El Código Civil español parece seguir este criterio en su artículo 333, en el que dispone que "[t]odas las cosas que son o pueden ser objeto de apropiación se consideran como bienes muebles o inmuebles".

Así, la calificación de "cosa", aunque jurídicamente correcta, resultaría simplista o reduccionista a la luz de lo que la IA es, puede hacer entre nosotros y el lugar que está ocupando en nuestro desarrollo en diversas áreas (si no en todas). En ese sentido, se debe considerar que la inteligencia de esta "cosa", que es su rasgo más distintivo, puede ser tomada como argumento para referirnos a ella de una manera diferenciada, mediante una denominación particular que así lo destaque. En ese sentido, vale la pena recordar que ya el Derecho ha mostrado ser flexible a variar sus referencias a ciertas realidades que antes denominaba "cosas", tomando en cuenta sus características más destacadas en nuestro relacionamiento con ellas. Así sucedió recientemente con los animales, a quienes se les ha denominado legalmente "seres sensibles" desde la modificación introducida al artículo 655 del Código Civil colombiano.[3] Por la misma vía, las capacidades de la IA, sus efectos y su potencialidad de implementación pueden generar un efecto similar, de manera que podamos variar la calificación que le damos, en aras a considerar de manera más adecuada esas características y sus efectos en nuestro relacionamiento con ella.

Un punto adicional a favor de esta misma variación en la denominación está en la posibilidad de que, eventualmente, la IA sea capaz de ser consciente de sí misma. Ello representa que, en un futuro (que no parece muy distante) el debate respecto de la personalidad de estas formas de inteligencia se hará más álgido, en la medida en que ella se implemente en cuerpos humanoides, o seamos capaces de sentir empatía o emociones reactivas hacia ellas y sus características, o que lleguemos a afirmar la necesidad de que se ajusten también a deberes jurídicos desde sus propias capacidades como artefactos inteligentes.[4] En esa medida, ya se ha discutido la posibilidad de predicar de las máquinas una personalidad electrónica (*"electronic personhood"*), sin que hasta el momento haya una tendencia académica al respecto en algún sentido

3 Tal vez, el mejor ejemplo de esta flexibilidad aparece en punto de la abolición de la esclavitud. En particular, porque en la época romana se consideró al esclavo como una cosa corpórea. Así lo registra el jurista Gayo (Inst. 2, 12) al hablar de la clasificación de las "cosas" corporales e incluir en ella a los esclavos.

4 Al respecto, la filosofía moral ha encontrado en las actitudes reactivas una base para la responsabilización interpersonal y la construcción de la moralidad. Sobre este punto, Strawson (1962) y Darwall (2006).

determinado (sobre este punto: Ávila, 2021). Por supuesto, todavía se puede debatir si es posible siquiera hablar de una verdadera autonomía o inteligencia en estas creaciones, pero no por ello podemos ignorar que, si tienen una real capacidad de razonamiento lógico, la consciencia de sí mismas no puede ser tomada como una imposibilidad radical si ellas son potenciadas en su independencia. De hecho, algunos reportes periodísticos ya han dado cuenta de ese eventual escenario, lo que descarta que se trate de una imposibilidad radical.[5]

En esa línea, la designación de la IA como "cosa", tomada como "una contraposición a la persona o sujeto de derecho" y, por ende, "el objeto de las relaciones jurídicas" (Arévalo, 2022: 129), quedaría en entredicho, como quiera que la interacción con ella revelaría que no se trata de una realidad meramente pasiva en nuestras dinámicas actuales, como tampoco algo que no tenga posibilidades de personificación ni potencialidad para fungir como sujeto racional y, eventualmente, consciente. Esa eventual consciencia, por supuesto, también pondría en duda la "apropiabilidad" que de las cosas se predica para que sean designadas como bienes por parte del Derecho.

En ese sentido, el Derecho de la propiedad intelectual podría darnos mejores herramientas para una denominación más ajustada a la realidad, pues en él las cosas producto del intelecto se les denomina "creaciones intelectuales" (Guerrero, 2023). Esta denominación no riñe con otras que también son pertinentes en la actualidad para efectos de la regulación de esta nueva realidad, como sucede con el término "producto", que refiere a algo que previamente tuvo un proceso de elaboración y que, después de finalizado, ha sido puesto a disposición de otros para su uso o disfrute dentro del mercado. En efecto, la IA se trata también de un producto que es una creación intelectual.[6] Sin embargo, esta expresión

---

5 Francesc Bracero, "Así convence un robot a otro para rebelarse", *La Vanguardia*, 27 de noviembre de 2024. Disponible en https://www.lavanguardia.com/vida/20241127/10144165/asi-convence-robot-rebelarse.html. También, "Informante: IA 'sensible' de Google contrató a un abogado", *DW*, 1 de julio de 2022. Disponible en https://www.dw.com/es/inteligencia-artificial-sintiente-de-google-contrató-a-un-abogado-asegura-ingeniero-suspendido/a-62335042.

6 Así se desprendería de la definición que trae el artículo 4, numeral 1 de la Directiva 2024/2853 del Parlamento Europeo y del Consejo, sobre responsabilidad por los daños causados por productos defectuosos.

también podría quedarse "a mitad de camino", por el enfoque que pone en el intelecto del creador y no en el de la cosa o producto creado. Por lo mismo, podría complementarse la denominación para que las IA sean tenidas como creaciones intelectuales "artificialmente inteligentes" (CIAI), denominación que permitiría diferenciarlas de otras creaciones y separarlas de la muy general denominación de "cosas", además de ponernos en la tarea de reflexionar respecto de las implicaciones de su inteligencia. Es más, dentro de esta categoría podría especificarse que habrá unas IA con mayores capacidades de autonomía que otras, las cuales podrían denominarse, eventualmente, "agentes artificiales autónomos" (AAA), cuando se trate de IAs que actúan de manera mucho más independiente y respecto de las cuales no tenemos propiamente relaciones de subordinación sino de coordinación.

¿"Peligrosa"?

Ahora bien, en lo que tiene que ver con la peligrosidad de la IA, también parece haber una insuficiencia de los términos con que tradicionalmente entendemos que algo es peligroso. Para entender por qué, examinemos la noción de "peligrosidad" implementada por la doctrina y la jurisprudencia nacionales.

En nuestra tradición de Derecho civil continental, se evidencia que parte de la doctrina ha distinguido entre la peligrosidad de las actividades y la peligrosidad en la estructura de las cosas, distinción que ha sido acogida en el ámbito colombiano para resolver problemáticas asociadas a la responsabilidad civil derivada de escenarios que se reputan peligrosos (al respecto: M'Causland, 2020: 94). En virtud de esta distinción, aquello que denominamos peligroso lo puede ser, o bien en virtud de la manera como se realiza una actividad (ej. Implementando velocidad o fuerza), o bien por la estructura de una cosa que tiene un alto potencial dañino en sí misma considerada (ej. Explosivos, radiación, energía nuclear, etc.). Esta distinción conlleva preguntarnos sobre su pertinencia respecto de las IA o su uso.

Tratándose de la peligrosidad de la actividad, habría que preguntarnos sobre la posibilidad misma de que dicho criterio sea implementado

ante esta nueva realidad. Si las IA son una "cosa", o una creación intelectual, entonces, ¿de qué actividad relacionada con ella podríamos predicar tal peligrosidad? La discusión puede ser mucho más profunda, porque implicaría tomar partido ante la pregunta sobre si algo puede ser peligroso en sí mismo o no, indagación que se complica si se tiene en cuenta que una IA "peligrosa" probablemente sería una que ha sido creada con bajos estándares de seguridad para evitar daños a los usuarios, a menos que esa peligrosidad se haya derivado exclusivamente de las capacidades de aprendizaje autónomas de la IA y escapen al control del humano creador. Sobre tal cuestión volveremos más adelante.

Por el momento, y en relación con las actividades que se pueden desplegar respecto de la IA por parte de los seres humanos, diría que es posible distinguir, al menos, dos tipos. Las primeras son las actividades de tipo activo, que conciernen a la construcción, elaboración, programación o edición de la IA; las segundas, son las de tipo pasivo o de usuario, que tienen que ver con la interacción o implementación que hace de la IA, como producto finalizado, cualquier persona.

Respecto de las primeras, parece descabellado afirmar que programar un *software* se trate de una actividad peligrosa, al menos así lo demuestra la experiencia en relación con tantos otros programas que han sido elaborados y que no necesariamente dejan efectos nocivos en quienes trabajan en ellos. El creador despliega su intelecto, sin que normalmente ello repercuta en daños para sí o para terceros por el hecho de realizar esa tarea.

Respecto de las segundas, aún no se tienen datos suficientes para concluir que el uso de un *software* inteligente constituya un acto peligroso, aunque ya haya noticias de algunos delirios provocados en ciertas personas que usaron *chatbots* impulsados con IA, o ciertos indicios de dependencia emocional que resultaron en autolesiones o muerte por parte de quienes interactuaron afectivamente con ellas.[7] En estos casos, la novedad de esta tecnología no permite advertir suficientemente los riesgos derivados de su uso, todavía, razón por la cual habrá que esperar a contar con más información para tal efecto. En el entretanto, en

7 Nadine Yousif, "Los padres de un adolescente que se quitó la vida demandan a OpenAI, creadora de ChatGPT", *BBC Mundo*, 27 de agosto de 2025. Disponible en https://www.bbc.com/mundo/articles/c30z5lyjzygo.

situaciones donde se comprometa la seguridad de personas o bienes especialmente protegidos, habría que considerar la posibilidad de actuar precautoriamente mediante la advertencia de los posibles riesgos asociados a la actividad.

Ahora, si fuera el caso que se evidenciara alguna "peligrosidad" en las IA, ¿en qué criterio estaría asentada? En el caso colombiano, el carácter peligroso de una actividad ha sido determinado por la jurisprudencia civil y de lo contencioso administrativo mediante varios criterios. Algunos de ellos resultarían insuficientes para el ámbito de la IA, porque aluden a características físicas que no se evidencian en las propiedades de un *software,* tales como "el desequilibrio de fuerzas", "el uso de cosas o energía que puede causar daños a terceros", las "labores que conllevan al empleo de máquinas o la generación, distribución o almacenamiento de energías",[8] "el empleo instrumentos, aparatos, energías o sustancias que ofrecen riesgos o peligros en razón de su instalación, de su propia naturaleza explosiva o inflamable o de otras causas análogas" (Valencia, 1998: 288), "la capacidad de destrozo que tienen sus elementos" (Tamayo, 1999: 322), el "desequilibrio o alteración en las fuerzas que de ordinario despliega una persona respecto de otra",[9] entre otras más. Todas estas características enfocadas en la energía o la fuerza física que una cosa puede desplegar en su entorno. Por lo que, bajo este criterio, la IA no sería peligrosa en sí misma, sino solo en la medida en que funcionara u operara dentro de estructuras capaces de generar tales desequilibrios en las fuerzas o detentar capacidad destructiva. En ese sentido, la implementación de IA en este tipo de circunstancias no haría variar el régimen de responsabilidad al que habitualmente estas estructuras están sometidas, y que suele ser objetivo. Tal sería el caso de infraestructura de servicios críticos, el uso aeronaves, la implementación de ciertos instrumentos médicos, el uso de armas de fuego o material explosivo, o la puesta en marcha de maquinaria pesada, entre otras.

---

8 Corte Suprema de Justicia de Colombia, Sala de Casación Civil y Agraria, sentencia del 25 de octubre de 1999, G. J. cclxi, pp. 874-885, M.P. Fernando Ramírez Gómez.
9 Corte Suprema de Justicia de Colombia, Sala de Casación Civil y Agraria, sentencia de 23 de octubre 2001, exp. 6315, M.P. Carlos Ignacio Jaramillo, reiterada en sentencia SC9788 de 29 de julio de 2015, M.P. Fernando Giraldo Gutiérrez.

Con todo, otros criterios también usados por la jurisprudencia y la doctrina podrían ser más favorables a la consideración de la IA como peligrosa. Ellos obedecen, principalmente, a la "imposibilidad de controlar o prever los efectos" (Tamayo, 1999: 322) de ella, "la posibilidad de que su ejercicio cause daños a terceros, aunque se desarrollen con suma prudencia",[10] lo "imprevisible de las consecuencias del daño",[11] entre otros similares. Bajo esa línea serán peligrosas las actividades y las cosas que, aunque se vigilen o se desarrollen de manera diligente o prudente, pueden llegar a generar efectos fuera del control que se ejerce sobre ellas, los cuales pueden resultar probables pero impredecibles en el tiempo, modo o lugar donde acontecerán, dada la capacidad que tenga de salirse del control del operador, vigilante o guardián de la cosa. Estas características podrían predicarse de una IA cuando quiera que ella sea de aquellas que puedan tomar decisiones que excedan el ámbito de control que tiene el operador humano y, en ese sentido, se entiendan fuera de sus posibilidades de predictibilidad, o le resulten ajenas a su control, por tratarse de decisiones tomadas a partir de las propias habilidades autónomas de la IA. Este, creo, podría ser el caso de los daños derivados del transporte por vehículos autónomos.

Aquí, se podría considerar que la relación entre la peligrosidad y la autonomía de la IA puede ser una que está en relación de proporcionalidad directa. En cambio, aquellas IA que se asemejen a una extensión de la voluntad y del intelecto humano, por obedecer a los comandos u órdenes que se le comunican, no revestirán, en principio, una peligrosidad en sí mismas, más allá de la peligrosidad de las actividades en las que se implementen. Su uso en tareas comunes será la regla y, por tanto, no habría razones para que la responsabilidad derivada de ellas sea agravada con un régimen objetivo, en principio. Esto no obsta para que pueda imponerse una responsabilidad de este tipo, no debido a la peligrosidad de la IA, sino bajo otro fundamento, como por ejemplo la protección del consumidor, o generar un incentivo para el desarrollo

10 Corte Suprema de Justicia de Colombia, Sala de Casación Civil y Agraria, sentencia de 18 de marzo de 1976, GJ CLII, n 2393, p. 73, reiterada por la misma corporación en la sentencia SC-5686 de 19 de diciembre de 2018, M.P. Margarita Cabello Blanco.

11 Corte Suprema de Justicia de Colombia, Sala de Casación Civil y Agraria, sentencia SC002 del 12 de enero de 2018, M.P. Ariel Salazar.

u oferta más seguros de estas tecnologías por parte de sus creadores o proveedores, respectivamente (sobre estos incentivos: Vladeck, 2014).[12] En este sentido, puede diferenciarse la responsabilidad entre personas consumidoras que hacen uso de IA como parte habitual de sus vidas y se causan daños, de la responsabilidad ante los consumidores que se presenta en relaciones asimétricas de contratación.

Por otro lado, IAs con alta capacidad generativa, capacidad de auto-programación, o posibilidades de regeneración o resistencia a la programación, o cualquier otro nivel de independencia o autopercepción que exceda la capacidad de control humano los resultados de su operación (porque no responden a los comandos especificados, por ejemplo), difícilmente podrían ser consideradas seguras o no-peligrosas. Aquí podrían incluirse aquellas IA que tienen capacidad de recodificarse, recodificar órdenes, interpretarlas nocivamente o ampliar autónomamente su rango de acción, que para el humano sería indeterminado o, en algún punto, impredecible. La propuesta pasa por considerar que ellas pueden ser tenidas como creaciones intelectuales que reportan un peligro respecto del cual algunas características tradicionales con las que jurídicamente se identifica la peligrosidad se quedan cortas, dado que no abarcan las posibilidades de descontrol que puede tener una creación intelectual según sus mismas capacidades de independencia que le da su carácter inteligente. En estos casos, los daños provenientes de este tipo de IA podrían conllevar la imposición de una responsabilidad objetiva, esta vez, asentada en su peligrosidad. Con todo, una IA que revista esta peligrosidad parece más una posibilidad en el futuro, pues aún no se conocen con certeza tecnologías inteligentes con tales características ni las probabilidades de que esto pueda suceder hoy en día. Este tipo de preocupación se relaciona, en particular, con lo que se ha llamado IA general (IAG) o IA fuerte, que es el nombre dado a una IA hipotética que, eventualmente, pueda realizar todas las tareas que realiza un ser humano y resolver los ilimitados problemas que se le

12 Al respecto, véase Directiva 2024/2853 del Parlamento Europeo y del Consejo, de 23 de octubre de 2024, sobre responsabilidad por los daños causados por productos defectuosos y por la que se deroga la Directiva 85/374/CEE del Consejo.

planteen.[13] Dicha IA, se teme, podría exceder las capacidades de control humanas y generar situaciones de daño.[14]

Pero, al considerar que la peligrosidad de la IA puede estar en función de sus capacidades de autonomía debe enfrentarse una cuestión importante que antes introducimos, y es que, en el contexto actual de regulación (escasa aún), existe un marco de *soft law* que ha dispuesto el deber de creación responsable y ética de las IA.[15] Esta es una razón para considerar que la creación de una IA que tenga capacidades de autonomía suficientes como para hacer daño a terceros resultará siendo una que será deficiente en sentido normativo, esto es, que no cumplirá con los estándares de seguridad que se desprenden del deber ético de que la IA no resulte nociva para las personas. Dicho con otras palabras, una IA que genera un daño sería una que pone en evidencia un trabajo deficiente por parte de su creador al momento de la programación. De alguna manera, el daño "hablaría por sí mismo" respecto de la deficiencia en la creación de esta inteligencia (una suerte de aplicación de la regla *res ipsa loquitur*). En esa medida, la responsabilidad por estos daños no sería objetiva realmente, pues en el fondo el daño se explicaría por una desviación conductual y subjetiva del programador o creador de la IA, respecto de su marco normativo de referencia, al no haber confeccionado la IA con la diligencia necesaria para evitar el daño. En ese sentido, la IA sería peligrosa en sí, pero por efecto de fallas en su programación, imputables al creador o desarrollador del *software*.

---

13 "¿Qué es una IA fuerte?", IBM. Disponible en https://www.ibm.com/mx-es/think/topics/strong-ai#:~:text=La%20inteligencia%20artificial%20fuerte%2C%20o,una%20gama%20ilimitada%20de%20problemas.

14 Desde la industria del entretenimiento, obras cinematográficas de ficción han imaginado esta posibilidad futura. Por ejemplo, las conocidas películas de *Terminator,* del director James Cameron (1984 y 1991) o, más recientemente, la serie *Murderbot* (2025) de Apple TV, dirigida y creada por Paul Weitz y Chris Weitz, recrean escenarios de una revolución violenta de las máquinas humanoides impulsadas por IA.

15 V.g. Unesco, "Recomendación sobre la ética de la IA". Disponible en https://unesdoc.unesco.org/ark:/48223/pf0000381137_spa. Última vez consultado el 22 de octubre de 2025.

Criterios de la reglamentación europea en la determinación del riesgo de la IA

Las anteriores reflexiones han sido tomadas en consideración a las ideas que se han expuesto en el ámbito de la responsabilidad civil, que como se ha visto, tienen que ver con criterios referidos, principalmente, a la capacidad de destrucción de las cosas o a la imprevisibilidad de que el riesgo se concrete o sea evitado por el ser humano, aunque actúe diligentemente. Estos criterios, como ha quedado en evidencia, pueden no resultar del todo satisfactorios respecto de la no corporalidad de la IA, o por no tener en cuenta que no siempre es inminente o previsible que ella pueda salirse del control humano si se considera sus capacidades de independencia.

Adicionalmente a esta circunstancia, el Reglamento de IA aprobado por la Unión Europea recientemente (UE 2024/1689), conocido como la "Ley de IA" acoge criterios sobre el riesgo de ella que no son del todo uniformes, ni tienen que ver necesariamente con la imprevisibilidad o la capacidad de destrucción física de las cosas a terceros, sino con el ámbito en el cual ella puede causar el daño o la importancia que le damos a ciertos derechos que pueden resultar afectados.

Para evidenciar esto, veamos cómo el Reglamento se refiere a los riesgos derivados de la IA, en particular, en sus artículos 5, 6 y 7, referentes a las prácticas no permitidas, las IA de alto riesgo y los criterios para modificar o suprimir sistemas de IA del listado de IAs de alto riesgo, respectivamente.

Prácticas no permitidas

En primer lugar, el Reglamento prohíbe explícitamente ciertas prácticas o implementaciones de la IA consideradas sumamente perjudiciales e incorrectas, debido a que contravienen los valores fundamentales de la Unión Europea, como el respeto a la dignidad humana, la libertad, la igualdad, la democracia y el Estado de Derecho. La introducción en el mercado, la puesta en servicio o la utilización de los sistemas de IA relacionados con estas prácticas quedan prohibidas. Tales riesgos

prohibidos pueden sintetizarse en los siguientes:

• Uso de sistemas de IA que empleen técnicas subliminales o manipuladoras que alteren sustancialmente el comportamiento de una persona o colectivo, disminuyendo su capacidad para tomar decisiones informadas y llevándolos a actuar de manera diferente a como lo harían normalmente, con consecuencias potencialmente perjudiciales.

• Utilización de sistemas que exploten vulnerabilidades relacionadas con la edad, discapacidad o situación social o económica de personas o colectivos, con la intención o el efecto de modificar de forma significativa su comportamiento, generando posibles daños.

• Implementación de sistemas que evalúen o clasifiquen a personas según su comportamiento social o características personales reales o inferidas, cuando esto resulte en tratos desfavorables en contextos no relacionados con los datos originales, o en consecuencias injustificadas o desproporcionadas respecto al comportamiento evaluado.

• Empleo de sistemas destinados a valorar o predecir el riesgo de que una persona cometa un delito basándose únicamente en perfiles o características de personalidad, sin considerar hechos objetivos y verificables directamente relacionados con una actividad delictiva.

• Creación o ampliación de bases de datos de reconocimiento facial mediante la recolección no selectiva de imágenes obtenidas de internet o de sistemas de videovigilancia.

• Uso de sistemas diseñados para inferir emociones de personas en entornos laborales o educativos, salvo que estén destinados específicamente a fines médicos o de seguridad.

• Aplicación de sistemas de categorización biométrica que utilicen datos biométricos para deducir raza, opiniones políticas, afiliación sindical, creencias religiosas o filosóficas, vida sexual u orientación sexual, fuera de contextos legales justificados o de tratamiento de datos lícitos.

• Empleo de sistemas de identificación biométrica remota en tiempo real en espacios públicos con fines de cumplimiento legal, salvo en casos en que sea estrictamente necesario para alcanzar objetivos legítimos y específicos. Solo se permiten excepciones en situaciones muy limitadas y bajo autorización judicial o administrativa.[16]

Las aplicaciones anteriores de la IA están directamente prohibidas, pero no se encuentra en ellas un criterio similar al implementado en la doctrina de la responsabilidad civil, relacionado con el uso de fuerzas o imprevisibilidad del daño, sino, relacionado a la infracción de derechos reconocidos como fundamentales, como la privacidad, la igualdad, el *habeas data* y la libertad.

Sistemas de IA considerados de alto riesgo

En relación con otras implementaciones de la IA, el Reglamento entiende que se considerarán de alto riesgo los sistemas que cumplan alguno de los siguientes criterios:

Criterio 1. Integración o implementación de la IA a ámbitos previamente regulados según el Anexo I

Se considera de alto riesgo la IA cuando ella se integre como componente de seguridad de un producto o de una actividad cuya regulación debe ser armonizada según el Anexo I del Reglamento, o que se trate de un producto de ese tipo en sí mismo, y además deba someterse a una evaluación de la conformidad de terceros para su introducción en el mercado o puesta en servicio con arreglo a los actos legislativos de armonización de la Unión enumerados en aquel Anexo.

16 Se exceptúa de tal prohibición la búsqueda selectiva de víctimas concretas (secuestro, trata, explotación sexual, personas desaparecidas); la prevención de una amenaza específica, importante e inminente para la vida o la seguridad física de las personas o de un atentado terrorista real y previsible; la localización o identificación de una persona sospechosa de haber cometido un delito grave (castigado con una pena máxima de al menos cuatro años de privación de libertad).

En dicho Anexo I se enumera una amplia gama directivas que son objeto de armonización y que, por lo general, tienen que ver con productos y actividades que también se tienen como peligrosos dentro del ámbito de la responsabilidad civil. En efecto, se disponen dentro de estas actividades las siguientes: el uso o implementación de máquinas; la seguridad de los juguetes; las embarcaciones de recreo y a las motos acuáticas; los ascensores y componentes de seguridad para ascensores; aparatos y sistemas de protección para uso en atmósferas potencialmente explosivas; equipos radioeléctricos; equipos a presión; instalaciones de transporte por cable; equipos de protección individual; aparatos que queman combustibles gaseosos; productos sanitarios; productos sanitarios para diagnóstico *in vitro*; vehículos de dos o tres ruedas y los cuatriciclos; equipos marinos; equipos para interoperabilidad del sistema ferroviario; vehículos de motor y de sus remolques y equipos de aviación civil.

Es decir que, cuando quiera que la IA se implemente en una de esas actividades o productos como componente de seguridad, o se trate de uno de esos productos en sí misma (como sería el caso, por ejemplo de una máquina o nave totalmente autónoma que es conducida por IA), y además deba someterse a una evaluación de la conformidad de terceros para su introducción en el mercado o puesta en servicio (esto tiene que ver con los permisos de índole administrativa que se deben otorgar para esa puesta en el mercado), entonces la IA se considerará de alto riesgo.

Criterio 2. Ámbitos señalados en el Anexo III

El Reglamento también considera como de alto riesgo los sistemas de IA que se utilicen en una serie de ámbitos predefinidos que son sensibles por la naturaleza de los derechos que pueden verse comprometidos o por las condiciones de seguridad que se pueden ver afectadas. Tales ámbitos son los siguientes:

Datos biométricos. Incluye sistemas de identificación biométrica a distancia (excepto los de simple verificación de identidad), categorización biométrica basada en características sensibles o protegidas, y

reconocimiento de emociones. Estos sistemas implican riesgos particulares por tratar datos personales especialmente sensibles.

Infraestructuras críticas. Abarca sistemas de IA utilizados como componentes de seguridad en infraestructuras esenciales como redes digitales, tráfico, agua, gas, calefacción o electricidad. Su mal funcionamiento podría tener un impacto en la seguridad de las personas y en el aprovisionamiento de elementos esenciales para vivir dignamente.

Educación y formación profesional. Se refiere a sistemas utilizados para determinar el acceso o admisión a instituciones educativas, evaluar resultados del aprendizaje, definir el nivel educativo al que puede acceder una persona, y detectar comportamientos prohibidos durante exámenes. La IA en este ámbito puede afectar el futuro educativo y profesional de las personas.

Empleo y gestión laboral. Incluye sistemas usados en procesos de selección, contratación, publicación de ofertas, análisis de currículos, evaluación de desempeño, toma de decisiones sobre promociones o despidos, y supervisión del comportamiento laboral. Su uso puede influir directamente en los derechos laborales y en la igualdad de oportunidades.

Servicios esenciales públicos y privados. Cubre sistemas que evalúan el acceso a servicios públicos esenciales como asistencia social o sanitaria, la solvencia o puntuación crediticia de personas (excepto detección de fraude), evaluación de riesgos en seguros de vida y salud, y priorización de llamadas o recursos en situaciones de emergencia. Estos sistemas pueden condicionar el acceso a derechos básicos.

Fuerzas y cuerpos de seguridad. Comprende sistemas destinados a evaluar riesgos de victimización, usar polígrafos o herramientas similares, valorar la fiabilidad de pruebas, predecir reincidencia o probabilidad de delinquir (no exclusivamente por perfiles), y elaborar perfiles de personas durante investigaciones. Su uso debe estar basado en legislación aplicable y controlado para evitar abusos o discriminaciones.

Migración, asilo y control fronterizo. Incluye el uso de sistemas para detectar mentiras (como polígrafos), evaluar riesgos de seguridad o sanitarios de personas que cruzan fronteras, asistir en la evaluación de solicitudes de asilo o visado, y detectar o identificar personas físicas, excepto para verificar documentos de viaje. Estas aplicaciones requieren especial atención por su impacto en derechos fundamentales.

Justicia y procesos democráticos. Engloba sistemas utilizados por autoridades judiciales para interpretar hechos y normas o resolver conflictos, y aquellos diseñados para influir en el voto o en el resultado de elecciones o referendos. Se excluyen herramientas internas de organización de campañas. Estos usos pueden afectar la imparcialidad judicial y la integridad democrática.

La calificación de un sistema de IA como de alto riesgo tiene como efecto principal que este sistema quede sujeto a un conjunto estricto y armonizado de requisitos y obligaciones obligatorias dentro de la Unión Europea. El objetivo de estas normas es garantizar un nivel elevado y coherente de protección de la salud, la seguridad y los derechos fundamentales de las personas, así como prevenir y mitigar riesgos inaceptables. Deben cumplir requisitos específicos que se centran en la fiabilidad, la transparencia y la minimización de riesgos a lo largo de todo su ciclo de vida. Estos incluyen implementar un sistema de gestión del riesgo, elaborar y mantener actualizada la documentación técnica antes de la introducción en el mercado o puesta en servicio; permitir el registro automático de acontecimientos ("archivos de registro") a lo largo de todo su ciclo de vida para garantizar la trazabilidad y facilitar la vigilancia post venta y la supervisión del funcionamiento; diseñarse con un nivel de transparencia suficiente para que los responsables del despliegue puedan interpretar y usar correctamente los resultados; ir acompañados de instrucciones de uso que detallen sus características, capacidades, limitaciones, niveles de precisión, solidez, ciberseguridad, y cualquier circunstancia conocida que pueda generar riesgos; garantizar que los supervisores humanos entiendan sus capacidades y limitaciones, sean conscientes del "sesgo de automatización" (confianza excesiva), puedan interpretar correctamente los resultados y tengan la capacidad de intervenir o detener el sistema si es necesario; deben alcanzar

un nivel adecuado de precisión, solidez (resistencia a errores y fallos) y ciberseguridad, y funcionar de manera uniforme en esos sentidos durante todo su ciclo de vida; ser resistentes a intentos maliciosos de terceros no autorizados de alterar su uso o funcionamiento, entre otros.

Criterios de excepción, modificación y supresión de las IA de "alto riesgo" en el Reglamento

Criterios para exceptuar la calificación de "alto riesgo"

A pesar de estar contemplados en el Anexo III, el Reglamento dispone que un sistema de IA no se considerará de alto riesgo cuando se cumpla alguna de las siguientes condiciones:

- Está destinado a realizar una tarea de procedimiento limitada.
- Está destinado a mejorar el resultado de una actividad humana previamente realizada.
- Está destinado a detectar patrones de toma de decisiones o desviaciones respecto a patrones anteriores, pero no sustituye ni influye en la valoración humana sin una revisión adecuada.
- Está destinado a realizar una tarea preparatoria para una evaluación que sea pertinente a efectos de los casos de uso enumerados en el Anexo III.

Así, no se estimará de alto riesgo la IA que cumple cualquiera de las condiciones anteriores, y además, no planteen un riesgo importante de causar un perjuicio a la salud, la seguridad o a los derechos fundamentales. Es decir, que se trata de dos condiciones que deben operar conjuntamente, una general de no representar riesgo importante para la salud, seguridad o derechos fundamentales, y una específica de estar en alguna de las condiciones antes numeradas. Aunque vale la pena considerar que estas condiciones tienen que ver, de manera general, con que la IA no incida sustancialmente en las decisiones adoptadas por el operador o usuario humano. En todo caso, el Reglamento dispone que

siempre se considerarán de alto riesgo los sistemas de IA que efectúan la elaboración de perfiles de personas físicas.

Criterios de modificación del listado del Anexo III

El Reglamento no utiliza una lista estática de sistemas de alto riesgo. El Artículo 7 establece un mecanismo dinámico (ajustable en el tiempo) que permite a la Comisión actualizar la lista de alto riesgo mediante actos delegados. Esto es crucial para un campo en rápida evolución como la IA, donde la tecnología y los riesgos asociados cambian constantemente. Este enfoque garantiza que el Reglamento sea resiliente a los avances tecnológicos. Este artículo 7 dispone los criterios bajo los cuales es posible añadir o modificar casos de uso de sistemas de IA en el Anexo III (sobre sistemas de IA de alto riesgo) o suprimir esos sistemas de IA de alto riesgo del Anexo III. El propósito fundamental de este mecanismo es garantizar que el reglamento se mantenga actualizado y flexible frente a la rápida evolución tecnológica.

En ese orden de ideas, se puede añadir o modificar casos de uso de sistemas de IA de alto riesgo en el Anexo III mediante actos delegados, siempre que se cumplan dos condiciones acumulativas:

La primera tiene que ver con el ámbito de uso. Los sistemas de IA deben estar destinados a ser utilizados en cualquiera de los ámbitos ya enumerados en el Anexo III (por ejemplo, biometría, infraestructuras críticas, educación, empleo, etc.).

La segunda se relaciona con el nivel de riesgo. Los sistemas de IA deben plantear un riesgo de perjudicar la salud y la seguridad o de tener repercusiones negativas en los derechos fundamentales. Este riesgo debe ser equivalente o mayor que el riesgo de perjuicio o de repercusiones negativas que plantean los sistemas de IA de alto riesgo que ya se mencionan en el Anexo III. Ahora bien, para determinar si se cumple la condición del nivel de riesgo, se deben considerar exhaustivamente una serie de criterios que a continuación se sintetizan:

- **Finalidad prevista.** La finalidad específica para la que se destina el sistema de IA.

- **Medida de uso.** La extensión en que el sistema de IA ha sido o es probable que sea utilizado.

- **Datos tratados.** La naturaleza y la cantidad de los datos tratados y utilizados por el sistema de IA, en particular si se incluyen categorías especiales de datos personales.

- **Grado de autonomía.** El nivel de autonomía del sistema y la posibilidad de que un ser humano anule sus decisiones o recomendaciones que puedan dar lugar a un perjuicio.

- **Historial de daños.** Si el uso del sistema ya ha causado perjuicio a la salud, seguridad o derechos fundamentales, o si existen problemas significativos y documentados relacionados con la probabilidad de tales perjuicios.

- **Alcance del perjuicio.** El posible alcance de las repercusiones negativas, considerando su intensidad y su capacidad para afectar a varias personas o afectar de manera desproporcionada a un colectivo específico.

- **Dependencia del resultado.** La medida en que las personas que podrían sufrir el perjuicio dependen del resultado generado por el sistema, especialmente si no es razonablemente posible renunciar a dicho resultado por motivos prácticos o jurídicos.

- **Desequilibrio de poder.** La medida en que existe un desequilibrio de poder o si las personas que podrían sufrir perjuicios se encuentran en una posición de vulnerabilidad respecto del responsable del despliegue (debido a su situación, autoridad, edad, etc.).

- **Facilidad de corrección.** La facilidad con la que se puede corregir o revertir el resultado generado, teniendo en cuenta que los resultados que afectan negativamente a la salud, la seguridad o los derechos fundamentales no deben considerarse fáciles de corregir o revertir.

- **Beneficio del despliegue.** La probabilidad y la magnitud del beneficio que el despliegue del sistema resultaría para las personas, los

colectivos o la sociedad en general, incluidas posibles mejoras en la seguridad de los productos.

• **Derecho de la Unión vigente.** Si el Derecho de la Unión ya establece vías de recurso o medidas efectivas para prevenir o reducir notablemente los riesgos planteados por el sistema (excluyendo acciones por daños y perjuicios).

Criterios de supresión o eliminación de la IA como una de alto riesgo

Finalmente, según el Reglamento es posible que se adopten actos delegados para suprimir sistemas de IA de alto riesgo de la lista del Anexo III, siempre que se cumplan dos condiciones:

• Cese de riesgo importante. Los sistemas de IA de alto riesgo en cuestión ya no deben plantear riesgos considerables para los derechos fundamentales, la salud o la seguridad, teniendo en cuenta los criterios de evaluación de riesgo que se abordaron inmediatamente atrás.

• Nivel general de protección. La supresión no debe reducir el nivel general de protección de la salud, la seguridad y los derechos fundamentales con arreglo al Derecho de la Unión.

En síntesis, la combinación de ambos criterios implica que dejar de considerar una IA como de alto riesgo no puede, de ninguna manera, resultar en una desprotección respecto del nivel de seguridad garantizado en el estado actual de la reglamentación.

La regulación europea, ¿un modelo a seguir?

La exposición de los anteriores criterios contenidos en la reglamentación europea en torno a la designación de una IA como de alto riesgo, su prohibición como tal, su modificación o exención nos deja algunas enseñanzas que vale la pena señalar en relación con la manera como se aborda el riesgo en la reglamentación y su relación con la manera como la responsabilidad civil se refiere a él.

En primer lugar, a diferencia de la doctrina de la responsabilidad civil, se evidencia una mayor dispersión de criterios para determinar que la IA puede ser riesgosa, esto sucede porque, mientras en la doctrina de la responsabilidad civil los criterios de peligrosidad son, por lo general, la naturaleza destructiva del bien o la imposibilidad de controlar los efectos dañosos de la actividad o del bien actuando con diligencia, en el ámbito de la reglamentación aparece que la afectación a derechos fundamentales es una razón muy importante para considerar la IA riesgosa.

Así, mientras que solo algunos criterios implementados en el Reglamento tienen que ver con el poder de destrucción o desequilibrio en las fuerzas que pone en riesgo la vida, salud o seguridad de las personas en ciertos contextos,[17] y otros con la imposibilidad de controlar la IA mediante actuaciones diligentes,[18] muchos otros tienen que ver con la afectación de derechos fundamentales, tales como la privacidad, la igualdad, la dignidad, el debido proceso, el habeas data, o la misma idea de un balance democrático ante el Estado, entre otros.

Ello quiere decir que la afectación a los derechos fundamentales es un punto diferenciador entre ambos ámbitos, pues no está presente actualmente en la responsabilidad civil al momento de considerar una actividad o cosa como peligrosa. En particular, la responsabilidad civil suele echar mano de un criterio probabilístico en el que riesgo se considera como tal en función de la frecuencia del daño o de la magnitud de este, mientras que en el ámbito del Reglamento el carácter riesgoso de la IA tiene también que ver con que esta llegue a afectar derechos fundamentales, cuestión que ocupa buena parte de los criterios de riesgo implementados.

---

17 Así podría decirse de cuando el Reglamento establece que para la modificación de las actividades dispuestas en el Anexo III, se debe consultar el "historial de daños" de la actividad, o "el alcance del perjuicio", o cuando en el mismo Anexo se dispone como actividades donde la IA es de alto riesgo las relacionadas con "infraestructuras críticas" o el "acceso a servicios esenciales". También, el Anexo I dispone varias actividades típicamente riesgosas, como se anotó antes.

18 Así podría decirse de cuando el Reglamento establece que para la modificación de las actividades dispuestas en el Anexo III, que se acuda a ciertos criterios como el "grado de autonomía" de la IA, la "dependencia del resultado" y la "facilidad de corrección" del resultado o la operación realizada por la IA.

Claro, uno podría interpretar que una afectación a un derecho fundamental es una afectación muy grave que encaja en el razonamiento probabilístico, pero la diferencia está en que la valoración de un derecho fundamental no siempre es uniforme entre los participantes de una práctica cuando la afectación de ellos no necesariamente implica una afectación de la salud, la vida o la seguridad. El reglamento, en cambio, afirma que ciertas prácticas son riesgosas y prohibidas, aunque no necesariamente comporten lesiones a la salud, a la vida o a la seguridad de las personas, como sería el perfilamiento de ellas, o la obtención de datos biométricos, o la generación de sesgos, entre otros, que tienen que ver con asuntos democráticos de legitimidad y libertad, pero no necesariamente con daños o perjuicios psicofísicos. En esa medida, la sensibilidad o la proximidad del derecho fundamental con la intimidad, la igualdad, la dignidad o las libertades democráticas de las personas parece ser un criterio importante que opera como medidor del riesgo de la IA, lo cual tiene sus propios inconvenientes, en particular, porque la valoración de lo fundamental, cuando ella se independiza del daño psicofísico, es una cuestión debatida democráticamente que no siempre arroja resultados objetivos o que tiene una determinación objetiva entre los participantes. Esto se evidencia en que algunos de los criterios que el Reglamento usa para modificar el listado de actividades de alto riesgo del Anexo III (art. 7) parecen sumamente subjetivos o, al menos, requieren de mayor claridad para una comprensión uniforme o intersubjetivamente controlable.[19]

En ese orden de ideas, cabe reflexionar si es conveniente que la responsabilidad civil acoja estos criterios de riesgo para efectos de considerar una IA peligrosa o si, por el contrario, debe el Reglamento observar los criterios más tradicionales de peligrosidad de la responsabilidad civil, pero repensados para ser aplicados a las IA, es decir, tomando en cuenta su eventual capacidad de destrucción y la incapacidad humana de controlarlas. Ciertamente, esta es una discusión de política jurídica que tendrá que darse también en nuestro contexto. Por el momento, vale la pena considerar que el uso cada vez más extendido de la IA obliga a

19 Así puede decirse de los criterios que apelan a la "finalidad prevista" de la IA, la "medida de uso", el "desequilibrio de poder" que ella pueda generar, o el "beneficio del despliegue" de su operación.

pensar en la manera como ella pueda ser peligrosa dentro de los ámbitos en los que se está implementado y que, en un futuro próximo, será muchos. Por lo mismo, tal vez convendría que, dada la inmensa gama de derechos fundamentales y sensibles que pueden ser afectados por la IA, se mantengan los criterios tradicionales de peligrosidad que tienen que ver con el potencial destructivo y la incapacidad de controlar su operación, a fin de no terminar estableciendo que toda la IA puede ser "de alto riesgo". Mejor sería, tal vez, considerar esas implementaciones como "sospechosas" o "sensibles", y elevar los estándares de diligencia que se deben observar en ellos. Este puede ser el caso de las IA implementadas en los ámbitos relacionados con la educación, la selección laboral o la migración, por decir algunos. En ellos, o en otros, podría pensarse en imponer una responsabilidad objetiva, no necesariamente derivada de la peligrosidad de la IA, sino de la necesidad de protegernos de afectaciones en tales ámbitos y como forma de incentivar el desarrollo responsable de estos *softwares*.[20]

En segundo lugar, y relacionado con lo anterior, se puede afirmar que la regulación europea que hasta el momento se ha expedido constituye solamente un primer paso de lo que debería acontecer en los próximos años respecto de la regulación venidera, dado que es notorio que la ingeniería y el uso de las IA tienen aún un margen de crecimiento mucho más amplio respecto de su clasificación, límites, forma de operación, rastreo del razonamiento, conservación de datos, entre otras. De esta manera, en un futuro debería ser más fácil determinar si una IA es peligrosa según una clasificación que se implementara sobre su independencia, razonamiento autónomo, capacidades de enmienda, filtros de seguridad, posibilidades de revisión y control humanos, entre otros criterios que sería útil establecer dentro de una realidad muy amplia de tecnologías varias que se están desarrollando. Quiere decir esto que una parte importante de la regulación de la IA tendrá que hacerse cargo del componente técnico especializado de los programas que de manera vertiginosa se están produciendo hoy en día, a fin de poder regular de

20 Respecto de los diferentes fundamentos de la responsabilidad objetiva, véase Vargas (2023).

manera más organizada el ritmo de estos avances y las posibilidades de que de ellos se deriven perjuicios a terceros.

Entre tanto, parece conveniente implementar la responsabilidad civil como incentivo de una producción responsable de estas tecnologías, lo cual podrá lograrse mediante la aplicación de estándares de diligencia y, en los casos más sensibles, de responsabilidad objetiva, que no necesariamente se fundamentará en el riesgo o peligro de la IA, sino en la sensibilidad de los derechos que ella puede afectar y en la necesidad de proteger a las posibles víctimas de las afugias probatorias que podrían pasar al momento de probar la negligencia del demandado.

Conclusiones

La evolución tecnológica de la inteligencia artificial plantea retos regulatorios que van más allá de las categorías tradicionales de peligrosidad, exigiendo una reflexión profunda sobre la naturaleza y el alcance de los derechos que estas tecnologías pueden afectar. Resulta conveniente reconocer que no toda IA debe catalogarse automáticamente como "de alto riesgo"; en cambio, cabe identificar aquellos ámbitos sensibles y las implementaciones "sospechosas" que demandan mayores estándares de diligencia y, eventualmente, responsabilidad objetiva. Esta distinción permite centrar la atención en la protección efectiva de los derechos fundamentales, sin obstaculizar innecesariamente el desarrollo innovador.

Asimismo, la regulación debe seguir un criterio técnico especializado, capaz de adaptarse con agilidad a los avances vertiginosos de la inteligencia artificial, contemplando factores como la autonomía, la capacidad de auto-programación y la resistencia a controles externos. El marco normativo emergente debe generar incentivos para que los desarrolladores y proveedores asuman una responsabilidad activa, promoviendo la creación de sistemas confiables, transparentes y rastreables a lo largo de todo su ciclo de vida. Solo un enfoque de esta naturaleza garantizará una convivencia equilibrada entre la innovación tecnológica y la salvaguarda de los derechos.

Finalmente, la responsabilidad civil puede constituir un instrumento clave para orientar la producción y el uso responsable de la IA, sin que ello implique, necesariamente, aplicar regímenes de responsabilidad objetiva basados únicamente en la peligrosidad. La regulación debe atender a la realidad contextual de cada implementación, privilegiando la protección frente a potenciales daños sin caer en generalizaciones que puedan desincentivar el progreso. Este delicado equilibrio exigirá un diálogo constante entre las disciplinas jurídicas y técnicas, con la finalidad de construir un marco normativo coherente, justo y eficaz.

Referencias

Akubo, Aduku A., Okpanachi L. Odiji, and Muhammed B. Muhammed (2025). "Fifth Industrial Revolution and an Overview of its Impact on Human Resources". *Human Capital Analytics*: 65-80.

Arévalo, Ismael (2022). *Bienes. Constitucionalización del Derecho Civil*. 3 ed. Bogotá: Universidad Externado de Colombia.

Ávila, Sergio (2021). "Robot as Legal Person: Electronic Personhood in Robotics and Artificial Intelligence". *Frontiers in Robotics and AI,* 8, 789327. DOI: https://doi.org/10.3389/frobt.2021.789327.

Darwall, Stephen (2006). *The Second-Person Standpoint: Morality, Respect, and Accountability*. Cambridge: Harvard University Press.

Guerrero, Manuel (2023). *La propiedad industrial: teoría y práctica*. 1 ed. Bogotá: Universidad Externado de Colombia.

Iglesias, Juan (2010). *Derecho romano. Historia e instituciones*. 18 ed. Madrid: Sello Editorial.

Labañino, Lizbeth, Antonio Lorca, María De las Heras y Alejandro Campina (2025). "Evolución del concepto de inteligencia artificial en la literatura científica: Un análisis sistemático." *Digital Education Review*, 46: 65-76. DOI: https://doi.org/10.1344/der.2025.46.65-76.

M'Causland, María Cecilia (2020). *Responsabilidad civil por el ejercicio de actividades peligrosas en Colombia: balance reciente y aproximación crítica*. 1 ed. Bogotá: Universidad Externado de Colombia.

Perez-Ugena, María (2024). "La inteligencia artificial: definición, regulación y riesgos para los derechos fundamentales". *Estudios De Deusto,* 72 (1): 307-337. DOI: https://doi.org/https://doi.org/10.18543/ed.3108.

Planiol, Marcel y Georges Ripert (1997). *Derecho civil. Biblioteca Clásicos del Derecho Civil.* Vol. 3. México: Harla.

Samoili, Sofía, Montserrat López, Emilia Gómez, Giuditta De Prato, Fernando Martínez-Plumed y Blagoj Delipetrev (2020). *AI Watch. Defining Artificial Intelligence. Towards an operational definition and taxonomy of artificial intelligence.* Luxemburgo: Publications Office of the European Union. https://publications-jrc-ec-europa-eu.translate.goog/repository/handle/JRC118163?_x_tr_sl=en&_x_tr_tl=es&_x_tr_hl=es&_x_tr_pto=tc.

Strawson, Peter (1962). "Freedom and Resentment". En Gary Watson (editor), *Proceedings of the British Academy*. Oxford: Oxford University Press.

Tamayo, Javier (1999). *Tratado de responsabilidad civil.* Tomo II. Bogotá: Legis.

Ternera, Francisco (2014). *Bienes*. 1 ed. Bogotá: Universidad del Rosario.

Vargas, Alexander (2023). "La fundamentación plural de la responsabilidad objetiva. Un debate abierto". En E. Cortés y M. M'Causland (editores), *Responsabilidad objetiva: entre esquemas tradicionales y nuevas realidades.* Bogotá: Universidad Externado de Colombia, 77-128.

Valencia, Arturo (1998). *Derecho civil. De las obligaciones.* 20 ed. Bogotá: Temis.

Velásquez, Luis (2022). *Bienes*. 16 ed. Bogotá: Ibáñez.

Vladeck, David (2014). "Machines Without Principals: Liability Rules and Artificial Intelligence". *Washington Law Review,* 89: 117-150. Disponible en: https://digitalcommons.law.uw.edu/wlr/vol89/iss1/6

Sobre el autor

Alexander Vargas Tinoco es abogado de la Universidad Externado de Colombia. Doctor en Derecho, Economía y Empresa por la Universidad de Girona. Es docente investigador del Departamento de Derecho Civil de la Universidad Externado de Colombia.

alexander.vargas@uexternado.edu.co
https://orcid.org/0000-0002-3750-2800