¿Quién paga los daños que causa la IA? De la ética a la responsabilidad por productos defectuosos

Who Pays for the Damages Caused by AI? From Ethics to Product Liability

Autor

Íñigo Navarro Mendizábal
Universidad Pontificia Comillas
Email: inavarro@comillas.edu
https://orcid.org/0000-0002-8901-3611

Resumen

La IA ofrece grandes beneficios, *pero* también puede causar daños, lo que nos debe llevar a reflexionar ética y jurídicamente sobre ella. Desde la reflexión ética surge una conclusión evidente: la IA no debe causar daños y como reverso de lo anterior, si un sistema con IA causa daños habrá que fijar un responsable que indemnice a la víctima de los daños causados. La existencia de esta responsabilidad además fomentará la fabricación de unos sistemas de IA más seguros y confiables. Lamentablemente, la legislación vigente no basta para abordar estos problemas, por lo que hace falta la introducción de nuevas regulaciones. Desde la UE se están abordando estas cuestiones buscando mitigar la dificultad probatoria y estableciendo que el fabricante será el responsable de los productos que incorporan sistemas de IA y que no tienen la seguridad que cabría legítimamente esperar.

Abstract

AI offers great benefits, but it can also cause harm, which should lead us to think ethically and legally. From ethical reflection, one conclusion is obvious: AI should not cause harm, and, on the other hand, if an AI system does cause harm, it will be necessary to establish a responsible party to compensate the victim for the harm caused. This responsibility will also encourage the manufacturing of safer and more reliable AI systems. Unfortunately, existing legislation does not address these problems, and new regulations are needed. The EU is addressing these issues by seeking to mitigate the evidentiary difficulty, and establishing that the manufacturer will be responsible for products that incorporate AI systems that do not have the security that would legitimately be expected.

Key words

Inteligencia artificial; ética; responsabilidad civil; daños de la IA; productos defectuosos.

Artificial intelligence; ethics; liability; AI harms; defective products.

Fechas

Recibido: 22/02/2024. Aceptado: 28/05/2024

1. La IA: una reflexión adversativa

La tecnología es hoy nuestro contexto, nuestra circunstancia en términos orteguianos. Vivimos inmersos en ella y no podemos escapar de ella de ninguna manera. Por ejemplo, si damos un paseo por la montaña y nos quedamos sin cobertura, lo característico de ese momento es, precisamente, que estamos sin cobertura, es decir, seguimos viviendo en un mundo "con cobertura". Guste o no, nuestra vida está marcada por la tecnología.

Los contactos con las personas más queridas se realizan a través de redes sociales como WhatsApp, las escuelas se llenan de apps formativas y vídeos interactivos, las decisiones empresariales se adoptan teniendo en cuenta los datos procesados por algoritmos, decidimos a dónde ir de viaje con la ayuda de plataformas en línea y llevamos en nuestros bolsillos tecnología que no habría estado al alcance de las personas más multimillonarias de hace una década.

La IA aporta valor en casi todas las funciones humanas desde las más cotidianas, como es llevarnos a un destino sin perdernos y sin atascos, a dar a un paciente el tratamiento más eficiente al ser capaz de analizar una cantidad abrumadora de información y mejorar la investigación médica

La IA ha supuesto un avance más en este desarrollo tecnológico. La IA puede replicar y luego superar reacciones humanas ante problemas, conocer patrones para luego hacer predicciones, automatizar tareas complejas, dar sentido a los datos. La IA aporta valor en casi todas las funciones humanas desde las más cotidianas, como es llevarnos a un destino sin perdernos y sin atascos, a dar a un paciente el tratamiento más eficiente al ser capaz de analizar una cantidad abrumadora de información y mejorar la investigación médica (WHO, 2021; NTT Data, 2023).

Sin embargo, existe un aspecto tenebroso de la tecnología y de la IA. Es el que aparece en las series distópicas como *Black Mirror*[1] y en la realidad que poco a poco vamos viviendo. Existen riesgos por doquier. Las "herramientas digitales se están convirtiendo cada vez más en un instrumento de manipulación y abuso en manos de algunos agentes empresariales y también de gobiernos autocráticos cuyo objetivo es socavar los sistemas políticos democráticos, lo que podría conducir a un choque entre sistemas políticos; explica que el espionaje digital, el sabotaje, la guerra a baja escala y las campañas de desinformación suponen un desafío para las sociedades democráticas" (Parlamento Europeo, 2022, p. 4). Ya se habla del "capitalismo de vigilancia" en expresión de Zuboff, lo que exige una regulación para que no nos expropien "la experiencia humana como materia prima gratuita aprovechable para una serie de prácticas comerciales ocultas de extracción, predicción y ventas" (Zuboff, 2020, p. 9). Los procesos de selección de todo tipo, incluyendo los laborales, pueden sesgarse hasta el punto de socavar principios constitucionales o generar un paro tecnológico si no se logra una reconversión digital (Parlamento Europeo, La IA en la era digital, 2020, p. 77).

Sobre los daños que puede causar la IA en nosotros mismos cabe destacar el libro *Demencia digital. El peligro de las nuevas tecnologías* de Manfred Spitzer cuyas

1 Brooker, Charlie (2011-presente). *Black Mirror* [Serie de televisión]. Zeppotron/Endemol-Netflix.

conclusiones van en la línea de la necesidad de control y de reflexión, porque las nuevas tecnologías pueden causar daño (y, de hecho, según su dictamen, lo están causando). Sin embargo, el daño no es algo intrínseco a la tecnología, sino que depende del uso que se haga de ella, por lo que conviene regularla. Así sugiere, por ejemplo, aprender de otros peligros, como puede ser el alcohol, que solo se puede consumir a partir de cierta edad y está gravado con impuestos o los automóviles a cuyo acceso hay limitaciones, pensando incluso en que podría haber un "carnet de internet" o entrenar a los jóvenes en una "competencia mediática" (Spitzer, 2013, pp. 304-309).

El momento en el que estamos podría denominarse como el de la "reflexión adversativa" (Navarro Mendizabal, 2023, p. 919). Somos conscientes de las enormes ventajas de la IA, *pero…* y toda reflexión y consideración que se realiza siempre tiene ese pero… Da igual la faceta humana de la que tratemos, siempre vemos la ventaja de la IA y su *pero*.

2. Necesidad de una reflexión ética y jurídica

Es tarea nuestra establecer qué es una IA honesta, lo que conllevará prohibir la que no lo sea, la que presente unos riesgos inasumibles por nuestra sociedad y fijar normativa, controles y gestión y limitación de riesgos en los demás casos

¿Qué se debe hacer ante una nueva realidad como la IA que genera inmensos beneficios y a la vez tiene riesgos potenciales gravísimos? No se trata de prohibirla (lo que no tendría ningún sentido), creo que es imposible suspenderla temporalmente, y no nos debe llenar de miedos y congoja… se trata simplemente de elaborar una reflexión ética y jurídica que permita una innovación que sea favorable, que sea buena.

Esta tarea no es ajena a nuestra reflexión habitual. El jurista romano Ulpiano enunció las *tria iura praecepta*: *honeste vivere, alterum non laedere, suum cuique tribuere*; lo que significa: vivir honestamente, no dañar a otro y dar a cada uno lo que le corresponde (Ulpiano: Digesto 1, 1, 10, 1). Esta es nuestra tarea en los tiempos de la IA:

a. Establecer qué es una IA honesta, lo que conllevará prohibir la que no lo sea, la que presente unos riesgos inasumibles por nuestra sociedad y fijar normativa, controles y gestión y limitación de riesgos en los demás casos. La UE, a través de la Propuesta de la Ley de Inteligencia Artificial (LIA)[2] ha clasificado los sistemas con IA en:

2 La LIA ha sido aprobada por el Parlamento Europeo en primera lectura el 13 de marzo de 2024.
Propuesta de Reglamento del Parlamento Europeo y del Consejo por el que se establecen normas armonizadas en materia inteligencia artificial (Ley de Inteligencia Artificial) y se modifican determinados actos legislativos de la Unión
{SEC(2021) 167 Final} - {SWD(2021) 84 Final} - {SWD(2021) 85 Final}.
https://eur-lex.europa.eu/legal-content/ES/TXT/HTML/?uri=CELEX:52021PC0206

i. IA que presenta riesgos inaceptables y por ello debe ser prohibida (art. 5 LIA)[3].

ii. IA de alto riesgo que estará sometida a numerosos controles (arts. 6 a 29 LIA)[4]. Estaría dentro de este grupo, por ejemplo, la IA que pudiera incorporarse a un vehículo autónomo que, por el mero hecho de circular, ya presenta altos riesgos por su potencialidad de daño.

iii. IA con un riesgo limitado que estará sometida a diversas obligaciones siendo una de las más importantes la de transparencia (art. 52 LIA). En este nivel estaría la IA que se incorpora en un chatbot como los que en numerosas ocasiones nos aparecen en las páginas web, por ejemplo, para hacer el *check in* en un vuelo.

iv. IA con mínimo riesgo que no tendrá obligaciones especiales.

b. Evitar todo daño de los sistemas con IA y, en el caso de que se produzcan daños, tener un sistema eficaz de reparación y resarcimiento de los daños causados que tendrá que afrontar el responsable para con la víctima.

c. Dar a cada uno lo suyo y, como siempre, lo difícil es saber quién es cada uno y qué es lo suyo. Valga una idea para arrojar un poco de luz: en derecho, el enriquecimiento sin causa es una figura que sirve para solucionar los casos en los que el patrimonio de una persona se ve enriquecido a costa de otra que se empobrece, sin que haya una causa jurídica que legitime una atribución patrimonial de uno a otro. Si una empresa se enriquece utilizando los datos de una persona que se ve empobrecida porque le han sustraídos esos datos (que evidentemente tienen un valor de mercado, por cuanto le están sirviendo para enriquecerse a la empresa) y no hay causa jurídica para esa atribución patrimonial, resulta evidente que la empresa deberá compensar a la persona "dando a cada uno lo suyo".

Debemos evitar todo daño de los sistemas con IA y, en el caso de que se produzcan daños, tener un sistema eficaz de reparación y resarcimiento de los daños causados que tendrá que afrontar el responsable para con la víctima

Más allá de las *tria iura praecepta*, la ética, como parte de la filosofía moral que analiza la conducta humana y lo correcto

3 En el art. 5 LIA se prohíben varios tipos de IA por ser contrarias a los derechos fundamentales o la seguridad de las personas, tales como: el uso de técnicas subliminales, el aprovechamiento de vulnerabilidades de grupos específicos, la evaluación o clasificación de la fiabilidad de las personas basada en su conducta social o personalidad, y el uso de sistemas de identificación biométrica remota en espacios públicos con fines de aplicación de la ley.
El uso de sistemas de identificación biométrica remota en espacios públicos con fines de aplicación de la ley solo se permitirá en casos excepcionales y cuando sea estrictamente necesario para alcanzar uno de los tres objetivos siguientes: la búsqueda de posibles víctimas concretas de un delito, la prevención de una amenaza específica, importante e inminente para la vida o la seguridad de las personas, o la detección, localización, identificación o enjuiciamiento de la persona que ha cometido o se sospecha que ha cometido un delito grave. En todo caso será necesaria la autorización previa y la supervisión del uso de sistemas de identificación biométrica remota.

4 Sintéticamente, la LIA clasifica los sistemas de IA de alto riesgo en varias categorías, que incluyen (arts. 6 y 7 LIA):
- IA en productos regulados: como dispositivos médicos y juguetes.
- Infraestructuras críticas: como transporte y energía, donde fallos podrían tener impactos significativos.
- Educación y formación profesional: que determinan el acceso a la educación.
- Empleo, gestión de trabajadores y acceso a empleo.
- Servicios esenciales privados y públicos: como la administración de justicia y servicios de emergencia.
- Aplicaciones biométricas: identificación y categorización de personas.
Estos sistemas deben cumplir con estrictos requisitos de transparencia, seguridad y responsabilidad (arts. 8 a 29 LIA).

e incorrecto, es trascendental para reflexionar sobre la IA. De hecho, siendo la tecnología un campo en transformación y que está actualmente avanzando a una gran velocidad es fundamental que se analice éticamente para asegurar que el avance y la transformación se realizan correctamente. La innovación por la innovación, seguir adelante hasta donde se pueda, sin pararse ni un segundo a pensar qué es lo correcto y, sobre todo, qué es claramente incorrecto nos puede llevar a situaciones sin salida que después costará mucho desandar, como el gato que avanza por la rama hasta llegar al punto que se queda atrapado porque no puede avanzar, ni sabe volver.

3. Los principios éticos de la IA y el no dañar

A pesar de toda la amplitud señalada en cualquier reflexión ética sobre la IA debe aparecer un mínimo común: no debe dañar, el clásico *alterum non laedere* latino

La ética aplicada a la IA tiene infinitud de áreas que abordar y miles de problemas que resolver, que pueden ir desde los daños tanto individuales como sociales causados por la tecnología y la IA, hasta el debate sobre cuál es el estatus moral de los sistemas de IA o si llegará a haber una IA con autoconciencia y que sienta y cómo deberíamos tratarla. Además, el análisis ético es muy amplio y abarca desde las conductas humanas más directas que diseñan, construyen o usan la IA hasta el comportamiento de un sistema de IA en sí que opera con más o menos autonomía del usuario o del diseñador o fabricante del sistema.

Por otro lado, en la medida en que el comportamiento de un sistema de IA sea más impredecible los problemas éticos aumentarán y hay que repensar la responsabilidad (DIGNUM, p. 1) al tiempo que son diferentes los problemas que plantean la IA específica y la IA general[5].

A pesar de toda la amplitud señalada en cualquier reflexión ética sobre la IA debe aparecer un mínimo común: no debe dañar, el clásico *alterum non laedere* latino.

El no dañar es la esencia de las leyes de la robótica[6] que imaginó Isaac Asimov y que son citadas expresamente por las Normas de Derecho civil sobre robótica (Parlamento Europeo, 2017). Se trata de la más absoluta interdicción del daño a una persona, por encima de las instrucciones que un robot reciba. Dichas Normas de Derecho civil abogan por un marco ético claro, estricto y eficiente, destacando que se debe tener en cuenta "la seguridad y la salud humanas; la libertad, la intimidad, la integridad y

5 Podría hablarse de diversas aproximaciones éticas, todas ellas necesarias (DIGNUM, p. 2):
- "Ethics by Design: the technical/algorithmic integration of ethical reasoning capabilities as part of the behaviour of artificial autonomous system;
- Ethics in Design: the regulatory and engineering methods that support the analysis and evaluation of the ethical implications of AI systems as these integrate or replace traditional social structures;
- Ethics for Design: the codes of conduct, standards and certification processes that ensure the integrity of developers and users as they research, design, construct, employ and manage artificial intelligent systems".

6 "1.ª Un robot no hará daño a un ser humano ni permitirá que, por inacción, este sufra daño.
2.ª Un robot obedecerá las órdenes que reciba de un ser humano, a no ser que las órdenes entren en conflicto con la primera ley.
3.ª Un robot protegerá su propia existencia en la medida en que dicha protección no entre en conflicto con las leyes primera y segunda".

la dignidad; la autodeterminación y la no discriminación, y la protección de los datos personales" (Principio 10). Se trata de poner a la persona siempre en el centro, porque la tecnología debe estar al servicio de la persona.

Quizás el texto europeo clave sobre la ética de la IA son las *Directrices éticas para una IA fiable* realizadas por el grupo de expertos de alto nivel sobre inteligencia artificial y editadas por la Dirección General de Redes de Comunicación, Contenido y Tecnologías de la Comisión Europea (Comisión Europea, Directrices, 2019). Estas Directrices éticas son particularmente importantes porque sirven de base para todas las propuestas de regulación que la UE está haciendo sobre la IA. Véase el iter: primero análisis ético y, sobre él, la propuesta regulatoria (Declaración europea, 2023).

El no dañar es una exigencia clave y esencial para la IA, lo que contiene su reverso: si se daña, el responsable (que no debía haber dañado en primera instancia) está obligado a "desdañar", a indemnizar, dejar indemne a la víctima

La IA que se vislumbra en estas directrices éticas, está centrada en la persona, lo que supone que complementa sus capacidades y lo hace de tal manera que maximiza beneficios y minimiza riesgos (Comisión europea, Directrices, 2019, p. 5). En las directrices se comienza con la exigencia de respeto a los derechos fundamentales, pues no existe una legalidad *online* y otra *offline* y de la misma manera no hay una ética *online* que sea más relajada que la *offline*. Así, los sistemas de IA no pueden ser herramientas que vulneren o favorezcan la vulneración de los derechos fundamentales.

En una segunda capa que refiere a la ética de la IA, más allá del elemental respeto a los derechos humanos, uno de sus principios que se enuncia es el de prevención del daño (Comisión europea, Directrices, 2019, p. 15).

Tras los principios éticos, el grupo de expertos que elaboró las directrices éticas señaló los requisitos que debe tener la IA y entre ellos se encuentra la exigencia de rendición de cuentas (Comisión europea, 2019, requisito 7, p. 18) que evidentemente requiere que "Cuando se produzcan efectos adversos injustos, deberían preverse mecanismos accesibles que aseguren una compensación adecuada. El hecho de saber que se podrá obtener una reparación si las cosas no salen según lo previsto es crucial para garantizar la confianza. Se debería prestar atención a las personas o grupos vulnerables" (Comisión europea, 2019, p. 25).

El no dañar es una exigencia clave y esencial para la IA, lo que contiene su reverso: si se daña, el responsable (que no debía haber dañado en primera instancia) está obligado a "desdañar", a indemnizar, dejar indemne a la víctima.

4. La importancia de una regulación europea de la responsabilidad civil de la IA

De la exigencia ética de no dañar se deduce la necesidad de resarcir el daño y en eso consiste la responsabilidad civil: establecer una norma clara de quién y en qué circunstancias tiene que responder del daño causado. Si un vehículo autónomo atropella a un viandante, si un detector inteligente de fuego no funciona y se produce un incendio o si un robot

quirúrgico falla en plena operación, debe haber un responsable y una indemnización del daño causado. Lo contrario sería devastador para la confianza que podamos tener en la IA. En este contexto, la responsabilidad civil es una pieza clave porque garantiza que quien sufra un daño sea resarcido y reciba una indemnización promoviendo la innovación en sistemas IA que sean más eficientes al evitar la causación de daños.

Los daños jurídicamente relevantes que puede causar la IA son de muchos tipos "tanto materiales (para la seguridad y la salud de las personas, con consecuencias como la muerte, y menoscabos al patrimonio) como inmateriales (pérdida de privacidad, limitaciones del derecho de libertad de expresión, dignidad humana, discriminación en el acceso al empleo, etc.) y pueden estar vinculados a una gran variedad de riesgos" (Comisión Europea, Libro Blanco, 2020, p. 14).

El dilema entre innovación y protección de los ciudadanos frente a los daños debe quedar resuelto de manera nítida con el principio de reparación íntegra que gobierna todo el derecho de daños

La responsabilidad civil de la IA se encuentra en un difícil entrecruce de caminos. Por un lado sus funciones son, como mínimo, dos: "por un lado, garantizan que las víctimas de un daño causado por otros perciban una indemnización y, por otro, proporcionan incentivos económicos a la parte responsable para que no cause dicho perjuicio" y además está en el centro de una tensión jurídico económica: "deben garantizar siempre un equilibrio entre la protección de los ciudadanos frente a los daños y la posibilidad de que las empresas innoven" (Comisión Europea, Informe, 2020, p. 14).

El dilema entre innovación y protección de los ciudadanos frente a los daños debe quedar resuelto de manera nítida con el principio de reparación íntegra que gobierna todo el derecho de daños. Una innovación que no internalizara los daños que causa sería indeseable. Es una falacia decir que la reparación íntegra coarta el progreso tecnológico, al contrario, favorece un progreso tecnológico sostenible, aumenta la confianza en la tecnología y favorece la calidad de la innovación.

En la actualidad carecemos de una normativa europea específica sobre la responsabilidad civil de la IA y esta ausencia está generando problemas jurídicos, económicos, empresariales y de confianza en el mercado. El derecho de daños en este campo está fragmentado entre los diversos países de la UE (cfr. Comisión Europea, Liability for AI, 2019), lo que genera inseguridad jurídica a las empresas hasta el punto de que la evaluación de impacto de la Propuesta de Directiva sobre responsabilidad en materia de IA estimó que la adopción de medidas de armonización específicas "tendrían un impacto positivo del 5 al 7 % en el valor de producción del comercio transfronterizo pertinente en comparación con la hipótesis de referencia" (Exposición de Motivos de la Propuesta). Además, la inacción europea está llevando a que diversos países estén estudiando introducir normas estatales, lo que conllevaría una mayor fragmentación en una materia que, precisamente, se caracteriza por ser global. La valoración que hace la Exposición de motivos sobre la Propuesta de Directiva sobre responsabilidad en materia de IA resalta:

> En términos de impacto social, la Directiva aumentará la confianza de la sociedad en las tecnologías de IA y promoverá el acceso a un sistema judicial eficaz. Contri-

buirá a un régimen de responsabilidad civil eficiente, adaptado a las especificidades de la IA, en el que las demandas fundamentadas de indemnización por daños y perjuicios sean estimadas. El aumento de la confianza social también beneficiaría a todas las empresas de la cadena de valor de la IA, ya que el aumento de la confianza de los ciudadanos contribuirá a una adopción más rápida de la IA. Debido al efecto incentivador de las normas sobre responsabilidad, evitar las lagunas en materia de responsabilidad también beneficiaría indirectamente a todos los ciudadanos mediante un mayor nivel de protección de la salud y la seguridad (artículo 114, apartado 3, del TFUE) y la evitación de fuentes de riesgo para la salud (artículo 168, apartado 1, del TFUE).

Las normas que tenemos en la actualidad no bastan para regular la responsabilidad civil de la IA

El grave problema es que siendo cierto todo lo anterior, el legislador europeo sigue sin rematar la normativa.

5. Por qué no sirve el marco que tenemos para la responsabilidad civil de la IA

Las normas que tenemos en la actualidad no bastan para regular la responsabilidad civil de la IA (cfr. Comisión Europea, Informe sobre las repercusiones, 2020, pp. 6 y ss.). Las razones son muchas y a continuación veremos algunas de ellas que nos sirven para reflexionar sobre la complejidad que tiene el resarcimiento de daños y la atribución de responsabilidades en estos casos.

5.1. Por el desconocimiento de los riesgos

Todavía no sabemos del todo los riesgos a los que nos enfrentamos y la IA "podría generar riesgos para la salud mental de los usuarios, derivados, por ejemplo, de su colaboración con robots y sistemas con IA humanoide, en el hogar o en entornos de trabajo" (Comisión Europea, Informe sobre las repercusiones, 2020, p. 9). No sabemos si quiera si existen estos riesgos o si son menores o mayores y más extendidos a otros sistemas con IA.

5.2. Por la difícil prueba de la relación de causalidad

La relación de la causalidad es uno de los elementos necesarios para establecer la responsabilidad civil, porque hay que acreditar de qué acción u omisión trae causa el daño y esto no es fácil si hay una IA de por medio. Por ejemplo, puede haber una concurrencia de causas y ser muy difícil para la víctima averiguar de cuál deriva su daño, pues el error puede provenir de las fuentes de datos, de los algoritmos mal codificados, de la interpretación que ha hecho el usuario o la máquina, de sesgos sociales, culturales, lingüísticos, cognitivos… (Díaz Alabart, 2018, pp. 18 y 19).

Otro ejemplo: supongamos que una alarma que debe detectar los humos en un hogar inteligente no funciona porque hay un cable en mal estado o porque el firmware tiene algún error. Evidentemente es mucho más sencillo probar el cable defectuoso que el firmware defectuoso (Comisión Europea, *Liability for AI*, 2019, p. 20).

5.3. Por la complejidad, la conectividad y la dependencia de los datos

La complejidad de los sistemas es mucho mayor que en cualquier producto que tuviéramos antes. La IA, la IoT y la robótica se caracterizan por combinar "la conectividad, la autonomía y la dependencia de datos para llevar a cabo tareas con poco o ningún control o supervisión humanos" (Comisión Europea, Informe sobre las repercusiones, 2020, p. 2)

Si un sistema de IA actúa de manera autónoma, sin control o supervisión humanos inmediatos y con algoritmos de aprendizaje automático y autónomo, puede llegar a ser muy difícil o imposible de comprender, lo que configura el efecto "caja negra", por el que resulta casi imposible atribuir la responsabilidad a alguien

La complejidad alcanza al entorno entero en el que un sistema interactúa con otros dispositivos, productos, servicios… que configura un ecosistema complejo en el que el ecosistema puede llegar a ser tan inmenso como el propio Internet.

Puede ocurrir que el riesgo de un producto se derive de la conectividad del mismo, como por ejemplo:

- Una pulsera inteligente para niños que posibilita que el niño pueda ser rastreado y contactado (notificación RAPEX de Islandia publicada en el sitio web EU Safety Gate (A12/0157/19).
- Un vehículo cuyo software tiene fallos en la seguridad de tal modo que permite que un tercero no autorizado acceda al sistema de control (notificación RAPEX de Alemania publicada en el sitio web EU Safety Gate [A12/1671/15]).

5.4. Por la autonomía, la opacidad y el efecto "caja negra"

Si un sistema de IA actúa de manera autónoma, sin control o supervisión humanos inmediatos y con algoritmos de aprendizaje automático y autónomo, puede llegar a ser muy difícil o imposible de comprender, lo que configura el efecto "caja negra", por el que resulta casi imposible atribuir la responsabilidad a alguien.

5.5. Porque es muy difícil averiguar si una IA es diligente o culpable

La culpa es otro de los elementos que no siempre será fácil de definir en el contexto de la IA, pues siempre hace referencia a un estándar de actuación diligente. Pues bien, cuando nos encontramos con una IA que es innovadora, ¿con qué estándar podemos compararla?

5.6. Por las dificultades de la ciberseguridad, que es dinámica y afectará a toda la vida del sistema con IA

Probablemente no exista el código informático que podamos garantizar que ahora y siempre va a ser seguro frente ataques externos, salvo que no esté conectado a alguna red. En cuanto existe conectividad podemos enfrentarnos a un ataque actual o futuro y esto les ocurre a las redes más seguras del mundo, que periódicamente son ciberatacadas y a veces con éxito.

Los sistemas de IA deben garantizar la seguridad (incluida la ciberseguridad) no solo en el momento de la puesta en circulación, sino también durante la vida de los mismos a través de actualizaciones si fuera necesario

Fijémonos en una cuestión, cuando compramos una casa todos asumimos que la seguridad de la casa será cosa nuestra en el futuro y decidiremos si poner una alarma o no. En cambio, si compramos un teléfono móvil, sí que exigimos que el sistema operativo esté al abrigo de ciberataques… pero ¿cuánta seguridad es suficiente en materia de responsabilidad civil?

Los sistemas de IA pueden llegan a producir situaciones casi únicas en función de lo que vaya haciendo el usuario. Veamos nuestros propios teléfonos móviles: cada uno de nosotros ha descargado diferentes apps, que usamos de diversa manera y almacenamos datos distintos hasta el punto de que probablemente no hay dos teléfonos iguales.

Todo lo anterior debe tener en cuenta la dimensión temporal: los sistemas de IA deben garantizar la seguridad (incluida la ciberseguridad) no solo en el momento de la puesta en circulación, sino también durante la vida de los mismos a través de actualizaciones si fuera necesario.

6. Las propuestas de regulación de la UE

Para dar respuesta a las carencias de la situación actual, las propuestas de regulación de la UE sobre la responsabilidad civil de la IA se articulan a través de dos instrumentos diferenciados:

- Propuesta de Directiva relativa a la adaptación de las normas de responsabilidad civil extracontractual a la IA[7].
- Propuesta de Directiva del Parlamento Europeo y del Consejo sobre responsabilidad por los daños causados por productos defectuosos[8] (que incluirá muchos productos con IA).

7 https://eur-lex.europa.eu/legal-content/ES/TXT/HTML/?uri=CELEX:52022PC0496

8 https://eur-lex.europa.eu/legal-content/ES/TXT/?uri=CELEX:52022PC0495

6.1. Propuesta de Directiva relativa a la adaptación de las normas de responsabilidad civil extracontractual a la IA

Es necesario redefinir el concepto producto, de tal manera que se incluyen los productos de la economía digital tales como: el software o los productos que incluyen para su funcionamiento software, apps o servicios digitales como todos los productos inteligentes o los vehículos autónomos

Esta propuesta no pretende configurar un sistema de responsabilidad civil que se aplique a la IA, sino que se limita a regular y, para decirlo claro, facilitar la prueba por parte de la víctima en los casos en los que la responsabilidad civil derivada de los daños de un sistema de IA sea de carácter subjetivo, es decir, que sea con culpa. No configura un sistema de responsabilidad civil y de hecho en el art. 5 plantea que en la evaluación que se haga de la directiva deberá plantearse si es conveniente la implantación de un sistema de RC objetiva sin culpa.

Lo que se solucionará con esta propuesta son las dificultades probatorias antes señaladas mediante una obligación de exhibir la prueba por parte del demandado (que está en mejor posición para hacerlo que la víctima) y presunciones de culpa y causalidad en determinados casos.

6.2. Propuesta de Directiva del Parlamento Europeo y del Consejo sobre responsabilidad por los daños causados por productos defectuosos (ProDirRPD)

Con esta propuesta se busca incluir en la responsabilidad civil de productos defectuosos a aquellos productos que tengan sistemas de IA[9]. La aproximación es correcta, desde mi punto de vista, porque la IA no existe de forma aislada, sino que opera en un determinado producto. Para salvar los problemas técnico-jurídicos que tenemos en la actualidad, se redefine el concepto producto, de tal manera que se incluyen los productos de la economía digital tales como: el software o los productos que incluyen para su funcionamiento software, apps o servicios digitales como todos los productos inteligentes o los vehículos autónomos. Así dice el art. 4 ProDirRPD que producto es "cualquier bien mueble, aun cuando esté incorporado a otro bien mueble o a un bien inmueble; por 'producto' se entiende también la electricidad, los archivos de fabricación digital y los programas informáticos".

Por su parte un producto es defectuoso "cuando no ofrece la seguridad que el público en general tiene derecho a esperar, teniendo en cuenta todas las circunstancias" y se señalan algunas de las circunstancias que se deben tomar en consideración (art. 6. 1 ProDirRPD). Como en la regulación actual el defecto se mide a través del *Consumer expectation Test*.

La definición de las personas responsables es un poco más compleja, aunque queda clara. En primer lugar tenemos la responsabilidad del fabricante proclamada en el art. 7. 1. ProDirRPD que es "toda persona física o jurídica que desarrolla, fabrica o produce un producto o que manda diseñar o fabricar un producto, o que lo comercializa con

9 Esta propuesta de directiva ha sido aprobada en Primera lectura por el Parlamento Europeo el 12 de marzo de 2024.

su nombre o su marca, o que desarrolla, fabrica o produce un producto para su uso propio" (art. 4. 11 ProDirRPD). Así nos encontramos al fabricante total bien sea directo o indirecto mandando fabricarlo aunque no lo haga él mismo, al fabricante aparente porque lo comercializa con su nombre e incluye a quien no lo comercializa, sino que lo fabrica para uso propio (que también puede causar daños a terceros). Igualmente será responsable el fabricante de un componente defectuoso, cuando haya provocado que el producto sea defectuoso (art. 7. 1. pº 2º ProDirRPD).

Cuando el fabricante esté establecido fuera de la UE serán responsables el importador y el representante autorizado del fabricante definidos en el art. 4. 12 y 13 ProDirRPD. Si ninguno de los anteriores está establecido en la UE será considerado responsable el prestador de servicios de tramitación de pedidos a distancia, que se define el art. 4. 14 ProDirRPD como "toda persona física o jurídica que ofrezca, en el transcurso de su actividad comercial, al menos dos de los siguientes servicios: almacenar, embalar, dirigir y despachar un producto, sin tener la propiedad del producto en cuestión", con la excepción de los servicios postales, los servicios de paquetería y los servicios de transporte de mercancías.

Se considera fabricante a cualquier persona que realice una modificación sustancial (de acuerdo con la normativa nacional o europea) fuera del control del fabricante original, al que podríamos denominar el modificador

También se considera fabricante a cualquier persona que realice una modificación sustancial (de acuerdo con la normativa nacional o europea) fuera del control del fabricante original, al que podríamos denominar el modificador (art. 7. 4 ProDirRPD).

Siempre debe haber algún responsable y en el caso de que no pudiera identificarse al fabricante o a los operadores económicos mencionados, será responsable el distribuidor si se niega a identificar en el plazo de un mes al operador económico o a la persona que le suministró el producto (art. 7. 5 ProDirRPD). Esto también será de aplicación "a cualquier proveedor de una plataforma en línea que permita a los consumidores celebrar contratos a distancia con comerciantes y que no sea un fabricante, importador o distribuidor" (art. 7. 6 ProDirRPD).

Cuando existan diversos operadores económicos que sean responsables, lo serán conjunta y solidariamente frente a la víctima (art. 11 ProDirRPD).

En resumidas cuentas, con esta propuesta se solucionan de una vez por todas las dudas habituales sobre quién responderá si un coche autónomo causa daños o si el robot quirúrgico Da Vinci falla o qué ocurre con una app que usa un algoritmo cuyos resultados causan daños al usuario: si el producto no tiene la seguridad que cabría legítimamente esperar (lo que es probable si causa daños) responderá el fabricante (o figuras asimiladas mencionadas).

7. Conclusiones

(1) La IA ofrece grandes beneficios, *pero* también puede causar daños, lo que nos debe llevar a reflexionar ética y jurídicamente. (2) En la reflexión ética, un principio es evidente y está recogido en todas las declaraciones: la IA no debe causar daños. (3) Como reverso de lo anterior, si la IA causa daños, alguien tendrá que ser responsable e indemnizar los daños causados, lo que es de justicia y además fomenta una IA más segura y confiable. (4) Lamentablemente la normativa vigente es difícilmente aplicable en muchos casos de daños causados por la IA, por cuestiones probatorias y por la dificultad de averiguar quién es el responsable. (5) Como esta es una cuestión global mucho más que nacional, la UE está trabajando sobre propuestas para solucionar ambas cuestiones y que tengan un alcance europeo. (6) Destaca en la propuesta de la UE que se señala como responsable al fabricante de productos que incorporan sistemas de IA y que no tienen la seguridad que cabría legítimamente esperar. (7) Las propuestas suponen un gran avance sobre la legislación existente y debería aprobarse sin demora, porque en este momento los daños que causa la IA solo pueden reclamarse a través de una legislación que lo dificulta mucho.

Referencias

Artificial Intelligence Act (AIA). Propuesta de reglamento del Parlamento Europeo y del Consejo por el que se establecen normas organizadas en materia de inteligencia artificial (Ley de Inteligencia Artificial) y se modifican determinados actos legislativos de la Unión {SEC(2021) 167 final} - {SWD(2021) 84 final} - {SWD(2021) 85 final} https://eur-lex.europa.eu/legal-content/ES/TXT/HTML/?uri=CELEX:52021PC0206

Comisión Europea. (2020). Informe de la Comisión al Parlamento Europeo, al Consejo y al Comité Económico y Social Europeo, Informe sobre las repercusiones en materia de seguridad y responsabilidad civil de la inteligencia artificial, el internet de las cosas y la robótica COM/2020/64 final.

Comisión Europea. (2020). Informe de la Comisión al Parlamento Europeo, al Consejo y al Comité Económico y Social Europeo, *Informe sobre las repercusiones en materia de seguridad y responsabilidad civil de la inteligencia artificial, el internet de las cosas y la robótica* COM/2020/64 final.

Comisión Europea. (2020). Libro Blanco sobre la inteligencia artificial - un enfoque europeo orientado a la excelencia y la confianza, Bruselas, 19.2.2020 COM(2020) 65 final. https://eur-lex.europa.eu/legal-content/ES/TXT/PDF/?uri=CELEX:52020DC0065

Comisión Europea. Dirección General de Redes de Comunicación, Contenido y Tecnologías. (2019). *Directrices éticas para una IA fiable*. Oficina de Publicaciones. https://data.europa.eu/doi/10.2759/14078

Comisión Europea. Directorate-General for Justice and Consumers. (2019). *Liability for artificial intelligence and other emerging digital technologies*. Publications Office. https://data.europa.eu/doi/10.2838/573689

Comisión Europea. Expert Group on Liability and New Technologies New Technologies Formation. (2019). *Liability for AI and Other Emerging Digital Technologies.* https://www.europarl.europa.eu/meetdocs/2014_2019/plmrep/COMMITTEES/JURI/DV/2020/01-09/AI-report_EN.pdf

Declaración Europea sobre los Derechos y Principios Digitales para la Década Digital (2023/C 23/01) https://eur-lex.europa.eu/legal-content/ES/TXT/HTML/?uri=CELEX:32023C0123(01)

Díaz Alabart, Silvia. (2018). *Robots y responsabilidad civil*. Ed. Reus.

Digesto de Justiniano

Dignum, Virginia. (2018). Ethics in Artificial Intelligence: Introduction to the Special Issus. *Ethics and Information Technology, 20*, 1-3. https://doi.org/10.1007/s10676-018-9450-z

Navarro Mendizabal, Iñigo Alfonso. (2023). La responsabilidad civil de la IA en la Unión Europea. Dicen que USA diseña, China fabrica y Europa regula... En *Treinta años de la Unión Europea. Una visión desde el Derecho*. Tirant lo Blanch.

NTT Data. (2023, 19 de octubre). *Tecnología al servicio de la salud: el avance de la atención médica virtual en América y Europa*. MIT Technology Review. https://www.technologyreview.es/s/15840/tecnologia-al-servicio-de-la-salud-el-avance-de-la-atencion-medica-virtual-en-america-y

Parlamento Europeo. (2017). Resolución de 16 de febrero de 2017, con recomendaciones destinadas a la Comisión sobre normas de Derecho civil sobre robótica (2015/2103(INL)). https://www.europarl.europa.eu/doceo/document/TA-8-2017-0051_ES.html

Parlamento Europeo. (2022). Resolución de 3 de mayo de 2022, La IA en la era digital (2020/2266(INI)). https://www.europarl.europa.eu/doceo/document/TA-9-2022-0140_ES.html

Propuesta de Directiva del Parlamento Europeo y del Consejo relativa a la adaptación de las normas de responsabilidad civil extracontractual a la inteligencia artificial (Directiva sobre responsabilidad en materia de IA) COM/2022/496 final. https://eur-lex.europa.eu/legal-content/ES/TXT/HTML/?uri=CELEX:52022PC0496

Propuesta de Directiva del Parlamento Europeo y del Consejo sobre responsabilidad por los daños causados por productos defectuosos COM/2022/495 final. https://eur-lex.europa.eu/legal-content/ES/TXT/HTML/?uri=CELEX:52022PC0495

Propuesta de Reglamento del Parlamento Europeo y del Consejo por el que se establecen normas armonizadas en materia inteligencia artificial (Ley de Inteligencia Artificial) y se modifican determinados actos legislativos de la Unión {SEC(2021) 167 Final} - {SWD(2021) 84 Final} - {SWD(2021) 85 Final}. https://eur-lex.europa.eu/legal-content/ES/TXT/HTML/?uri=CELEX:52021PC0206

Spitzer, Manfred. (2013). *Demencia digital. El peligro de las nuevas tecnologías*. Ediciones B.

World Health Organization (2021). Global strategy on digital health 2020-2025. https://www.who.int/docs/default-source/documents/gs4dhdaa2a9f352b0445bafbc79ca799dce4d.pdf

Zuboff, Shoshana. (2020). *La era del capitalismo de vigilancia*. Ediciones Paidós.