1

¿PERO DE VERDAD HAY LEYES QUE NOS PROTEGEN DE LA INTELIGENCIA ARTIFICIAL?

1.1 EL "GOLPE DE CORBATA"

En el año 1993, hace ya 31 años, puse en marcha en Madrid, junto con otros socios, la primera empresa privada en España (y en Europa) que ofrecía al mercado servicios de traducción a cinco idiomas, basados en una plataforma de Inteligencia Artificial. La tecnología a su vez la adquirimos a una empresa norteamericana: Globalink Inc.

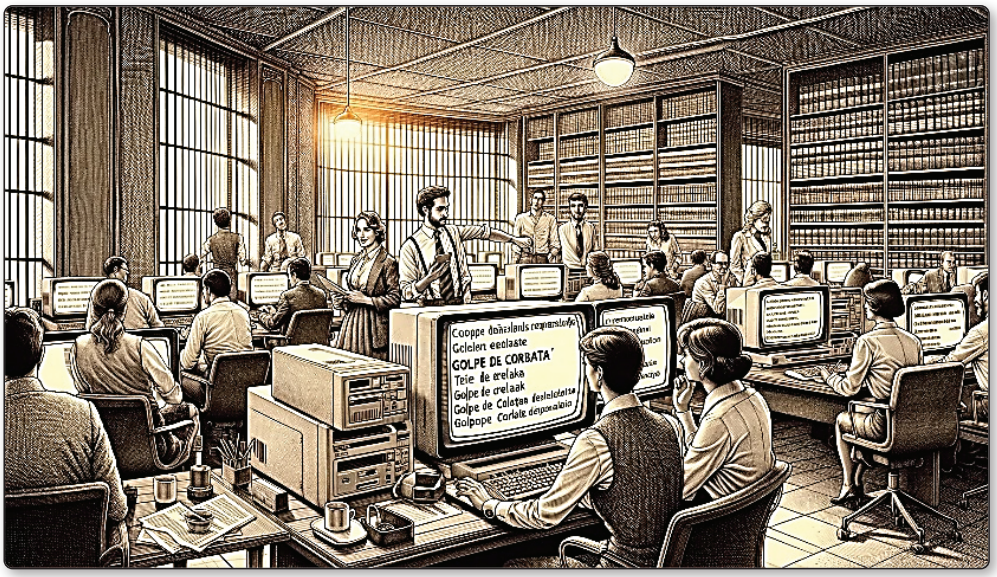
La empresa y su servicio de traducciones, comercializado en esos momentos por Telefónica Servicios Avanzados de Información, se llamaba HILO TRADUCTOR y se anticipó al uso de sistemas de IA en los procesos de traducción, ofreciendo servicios de traducción del español al inglés, francés, alemán, ruso y chino. Y pudiendo transferir los documentos traducidos a través de comunicaciones electrónicas. En esos momentos X 25. Salimos al mercado tres años antes de que empezase a comercializarse Internet en España. Todo un alarde.

Aunque el programa informático tenía una gran potencia, -traducía alrededor de 180.000 palabras a la hora-, le quedaba aún un buen trecho para mejorar en la calidad de las traducciones que producía, de tal forma que, una vez conseguida la traducción "en bruto" que había generado el programa, revisábamos cada documento con un equipo de traductores profesionales en esos idiomas, en los tres niveles fundamentales de revisión: el gramatical o semántico, el pragmático (es

decir, atendiendo al contexto o significado real que quiere expresar el autor y no solo al significado literal de las palabras) y si el texto lo requería por su carácter, el estilístico.

A la vez, ese proceso de revisión ponía en evidencia las deficiencias de la traducción generada por el programa, lo que llamábamos traducción en "bruto". Y esas deficiencias se identificaban y posteriormente se procuraban corregir ampliando las bases de datos de unidades semánticas a las que recurrían los algoritmos de traducción de programa.

Uno de nuestros clientes fue el Comité Olímpico, que nos solicitó la traducción del reglamento de las Paralimpíadas . Una tarde estaba junto con el equipo de revisores y escucho a uno de ellos que empieza a exclamar: ¡El golpe de corbata! ¡Qué ***** es el golpe de corbata! Efectivamente, nuestro querido algoritmo había traducido literalmente *"tie break"* como *"golpe de corbata"*, y no como desempate.



Más allá de las risas que nos provocó el suceso en su momento, la anécdota me sirve para transmitir las siguientes ideas:

- Toda tecnología puede fallar. Y la Inteligencia Artificial falla.
- La Inteligencia Artificial es un conjunto de tecnologías que están en un constante proceso de evolución. Y por lo tanto de ensayo, error y mejora.

- Es importante que las empresas, personas u organizaciones que diseñan modelos de Inteligencia Artificial, los entrenan y finalmente los ponen en explotación sean conscientes de esta circunstancia. Y que por lo tanto adviertan a sus usuarios de los riesgos que puede suponer su uso. Y en paralelo, establezcan controles para evitar un mal funcionamiento o corregir fallos que se detecten.

- Igualmente es importante que las personas, empresas o entidades que pretendan sacar al mercado servicios o productos en donde intervengan sistemas de Inteligencia Artificial sean lo suficientemente responsables para que ese sistema esté lo suficientemente probado a fin de evitar un pobre funcionamiento o potenciales perjuicios a sus usuarios.

Los ejemplos de legislaciones que vamos a ver a continuación en los siguientes apartados, tienen en gran medida como fin asegurar que esas tecnologías llegan al mercado con las mayores garantías, Lo iremos viendo a lo largo del capítulo.

1.2 EVITANDO UNA VISIÓN NEGATIVA Y ALARMISTA

Una idea clave que deseo destacar a partir de este momento es la de evitar abonarnos a una visión alarmista o negativa de la Inteligencia Artificial. Como todas las tecnologías tiene dos caras, y puede ser positiva o negativa en función del uso que hagamos de ella.

Nadie debe poner en duda que la_ lleva ofreciendo muchísimas ventajas y beneficios a la sociedad desde hace ya años, permitiendo abordar proyectos de todo tipo: de ingeniería, de medicina, de educación, de investigación biotecnológica, de comunicación, de ocio, de control y seguridad de infraestructuras críticas, de defensa, etcétera.

Y que, en una balanza, todas esas ventajas pesan enormemente más que los malos usos o los riesgos que pueda suponer.

Aun así, es deber de toda sociedad y sus legisladores regular el uso de esta tecnología, de la misma forma que se regula el uso de la telefonía o se ordena el tráfico aéreo. Y precisamente regularlo para que, con unas reglas claras y justas, podamos beneficiarnos de la misma y potenciar su desarrollo al máximo de sus posibilidades, con el fin de revertir esos logros técnicos en bienestar y riqueza para los ciudadanos.

1.3 UNA INQUIETUD LEGAL QUE VIENE DESDE HACE DÉCADAS

Una de las ideas que desde el principio debemos abandonar es la de pensar que los desarrollos tecnológicos que se introducen en las sociedades modernas se hacen de forma absolutamente descontrolada.

Sí es cierto que la velocidad a la que se están produciendo esos avances es cada vez mayor.

Hace cincuenta años los ciclos económicos asociados a las innovaciones tecnológicas abarcaban 10 a 20 años como mínimo, período en el que la tecnología se introducía, se expandía en su uso, se desarrollaba y llegaba a su período de madurez; para a su vez sustituirse o renovada al final de ese ciclo. La frecuencia de las oleadas de innovación tecnológica ha sido, paulatinamente, cada vez mayor, y al contrario, la duración de esos ciclos de crecimiento y sustitución de unas tecnologías por otras son cada vez más breves. Pasamos de los 20 o10 años de explotación de una tecnología a los 5, a los 3, y actualmente las novedades se suceden unas tras otras cada seis meses.

Aunque sí es cierto que muchas de esas innovaciones no son más que pequeñas variaciones de una tecnología o técnica de base; por ejemplo, el teléfono inteligente tal y como lo conocemos apareció con el primer IPhone en 2007, y a partir de ahí lo que se ha ido produciendo hasta nuestros días es una actualización de un concepto que ya tiene 14 años.

Otras tecnologías con las que vivimos diariamente, como por ejemplo el motor diésel, son ya centenarias. Lógicamente los actuales motores diésel son más eficientes que su prototipo de 1893, pero la esencia es la misma. La telefonía móvil que disfrutamos de forma extensiva desde hace 25 a 30 años tuvo su origen en 1973, hace ya 51 años. Las capacidades de conectividad y aplicaciones asociadas son innumerables; las mejoras en el rendimiento también, pero la esencia técnica es la misma.

Con estos ejemplos lo que quiero expresar es que vivimos dentro de un torbellino de información diaria sobre novedades tecnológicas que nos provocan la sensación de que la sociedad digital avanza a un ritmo desbocado, imposible de controlar, cuando, en la práctica, la base tecnológica y las funcionalidades principales de esas tecnologías tienen un tiempo de desarrollo mucho más amplio del que aparentemente pueda parecernos en función del bombardeo de información y publicidad diaria al que nos vemos sometidos.

Ahora bien, es indudable que la iniciativa innovadora y creativa de los seres humanos es impredecible. Por ejemplo: los ingenieros de Motorola que en 1973 hicieron la primera llamada desde un teléfono móvil no creo que se imaginasen a dónde ha llegado y evolucionado su invento.

Pero las líneas maestras de la finalidad o finalidades de cada tecnología, y los usos inicialmente previsibles que puede abarcar en sus primeros años de desarrollo, incluidos los que supongan converger con otras tecnologías complementarias, sí son predecibles.

Y por lo tanto, son predecibles los retos que esas tecnologías, sus potenciales aplicaciones y los previsibles desarrollos que puedan plantear a nuestra sociedad: los derechos ciudadanos, la seguridad jurídica de las personas, su impacto económico y social y las oportunidades que puede ofrecer a la prosperidad de las sociedades donde se introduzca.

En el caso concreto de la Inteligencia Artificial, la inquietud sobre el uso de esta tecnología por parte de los legisladores viene ya de lejos, desde el siglo XX, de la mano del procesamiento masivo de datos que supuso la extensión de la informática en los últimos 30 años de la pasada centuria, su tratamiento automatizado y los efectos que para los ciudadanos y sus derechos pudiera suponer un uso no consentido de sus datos personales o un tratamiento automatizado de dichos datos, que no hubiera sido claramente informado a los interesados. Y sobre todo, las consecuencias que para dichos ciudadanos pudieran suponer el resultado de dicho tratamiento automatizado.

Una de las primeras normas constitucionales que introdujo como derecho fundamental la protección de la privacidad de los ciudadanos frente al uso de la informática fue la constitución española de 1978 **en su artículo 18.4.**

> *"La Ley limitará el uso de la informática para garantizar el honor y la intimidad personal y familiar de los ciudadanos y el pleno ejercicio de sus derechos"*

Si avanzamos en el tiempo, de nuevo en España, la Ley Orgánica de Protección de Datos del año 1999, en su artículo 13, es de las primeras que ya define un derecho de los ciudadanos a ser informado e impugnar los resultados del uso de sus datos personales, con el denominado derecho de **Impugnación de Valoraciones:**

"Artículo 13. Impugnación de valoraciones

1. Los ciudadanos tienen derecho a no verse sometidos a una decisión con efectos jurídicos, sobre ellos o que les afecte de manera significativa, que se base únicamente en un tratamiento de datos destinados a evaluar determinados aspectos de su personalidad.

2. El afectado podrá impugnar los actos administrativos o decisiones privadas que impliquen una valoración de su comportamiento, cuyo único fundamento sea un tratamiento de datos de carácter personal que ofrezca una definición de sus características o personalidad.

3. En este caso, el afectado tendrá derecho a obtener información del responsable del fichero sobre los criterios de valoración y el programa utilizados en el tratamiento que sirvió para adoptar la decisión en que consistió el acto.

4. La valoración sobre el comportamiento de los ciudadanos, basada en un tratamiento de datos, únicamente podrá tener valor probatorio a petición del afectado."

La Ley Orgánica de Protección de Datos 1999 era producto de una transposición al derecho español de una Directiva de la Unión Europea del año 1995, la Directiva 95/46/CE, ya derogada, en donde se preveía precisamente este problema del uso de sistemas automatizados de tratamiento de datos y toma de decisiones y su impacto sobre los derechos ciudadanos, en concreto su artículo 15:

"Artículo 15. Decisiones individuales automatizadas

1. Los Estados miembros reconocerán a las personas el derecho a no verse sometidas a una decisión con efectos jurídicos sobre ellas o que les afecte de manera significativa, que se base únicamente en un tratamiento automatizado de datos destinado a evaluar determinados aspectos de su personalidad, como su rendimiento laboral, crédito, fiabilidad, conducta, etc.

2. Los Estados miembros permitirán, sin perjuicio de lo dispuesto en los demás artículos de la presente Directiva, que una persona pueda verse sometida a una de las decisiones contempladas en el apartado 1 cuando dicha decisión:

a) se haya adoptado en el marco de la celebración o ejecución de un contrato, siempre que la petición de celebración o ejecución del contrato presentada por el interesado se haya satisfecho o que existan medidas apropiadas, como la posibilidad de defender su punto de vista, para la salvaguardia de su interés legítimo; o

b) esté autorizada por una Ley que establezca medidas que garanticen el interés legítimo del interesado."

Como podemos ver, hace casi treinta años se previó el impacto que el uso de sistemas automatizados de toma de decisiones (algoritmos), que recurriendo a bases de datos masivas, -lo que ahora llamamos Big Data-, generasen a su vez perfiles de cada uno de nosotros en donde, de forma automatizada, valorasen aspectos que pudieran afectarnos en derechos ciudadanos como el acceso al empleo, la educación, la clasificación y categorización psicosocial o el acceso al crédito financiero.

La importancia que en el ámbito de la Unión Europea se dio a la regulación del uso de los datos personales y el aprendizaje institucional que la aplicación de las primeras normas de protección de los datos personales supuso, favoreció para que, a lo largo de este cuarto de siglo que ha pasado desde 1999, de la mano de la actualización y robustecimiento de las normas de protección de datos, el concepto de dato personal como base de la economía digital y su tratamiento masivo y automatizado a través del Big Data, fuera el fundamento para percibir la amenaza que la "economía del dato" pudiera suponer para los ciudadanos.

Con la aplicación del Reglamento General de Protección de Datos 679/2016 de la Unión Europea y otras legislaciones conexas se abrió la puerta para regular los más diversos aspectos de una sociedad digital en eclosión: desde la videovigilancia, pasando por las telecomunicaciones, la información sanitaria vinculada con los historiales de los pacientes, los servicios comerciales de la sociedad de la información o la ampliación del concepto de dato personal a un ámbito que abarca nuestra "huella", "rastro" o *"nube de datos"* que generamos a través de los dispositivos electrónicos que utilizamos, máquinas y electrodomésticos que usamos, geografía por la que nos desplazamos o páginas webs que visitamos, pasando por la imagen personal que es captada en los más diversos e insospechados lugares.

La creación a lo largo de estas dos últimas décadas de nuevos "derechos digitales", vinculados en muchos casos a salvaguardar a los ciudadanos y su privacidad del uso masivo de sistemas automatizados de búsqueda, como el **derecho al olvido**, sensibilizaron aún más a las autoridades de la necesidad de regular unas tecnologías que se estaban introduciendo de forma masiva en nuestra sociedad.

En el momento que estoy escribiendo estas líneas se ha aprobado definitivamente por el Parlamento Europeo el Reglamento Europeo (Ley) de la Inteligencia Artificial. Un hito legislativo a nivel mundial a la hora de definir los derechos ciudadanos frente al uso de esta tecnología y cómo se debe regular la misma para conseguir que sea ética, transparente y confiable.

Pero para llegar a esta meta hubo una serie de etapas.

La protección de datos fue la base, el hilo conductor, que ayudaba a definir las regulaciones de otras tecnologías conexas. La razón es sencilla. Un sistema de Inteligencia Artificial sin el uso de los datos es una cáscara vacía. No sirve para nada. La mitad de una tecnología de Inteligencia Artificial son los datos con los que se ha entrenado, y posteriormente los que usa para tomar las decisiones y aprender autónomamente en algunos de sus modelos.

Y cuando hablamos de datos hablamos de todo tipo de información. No ya los datos personales, sino también contenidos técnicos, artísticos, literarios, enciclopédicos, imágenes, registros electrónicos, direcciones IP, datos geográficos, datos registrales, estadísticas de todo tipo.

Los datos que consideremos más anodinos o extraños, como pudiera ser por ejemplo, la paleta de colores de un plátano/banana en función de su nivel de madurez: desde el verde hasta el marrón oscuro, pasando por el amarillo intenso, entre medias, puede suponer una información clave para un programa que identifique y categorice en tiempo real el nivel de madurez de la fruta y permita reducir el desperdicio alimentario.

Caso Práctico nº. 1: ¿Es la sonrisa un dato personal?

Hace seis años asistí a un congreso sobre "Amor e Inteligencia Artificial" que se celebró en Zaragoza, en mayo de 2018. Fue un congreso muy interesante y muy original. Una de las intervenciones que más me impresionaron consistió en un ponente que presentó un programa, en esos momentos experimental, que mediante óptica inteligente, reconocía los gestos del público asistente en el patio de butacas de la sala de conferencias, y categorizaba, uno a uno, en tiempo real, los gestos del público. De tal forma que, del resultado de los datos obtenidos, podía contabilizar cuántas veces habíamos aplaudido a cada uno de los ponentes, o cuántas veces habíamos sonreído por los comentarios de los conferenciantes, o si simplemente habíamos prestado atención, o por el contrario, nos habíamos dedicado a mirar la pantalla de nuestro móvil mientras la charla de turno se convertía en mero ruido de fondo.

Lo interesante y útil de esta aplicación, para los ponentes, es que la explotación de los datos obtenidos les ofrecía curvas de mayor atención e interés, minuto a minuto y las reacciones del público a cada uno de sus comentarios, gracias a esa óptica inteligente, aplicada en este caso a reconocer y categorizar emociones y actitudes humanas.

Lo peligroso es que ese análisis se podía llevar hasta la individualización absoluta. De tal forma que, pongamos un supuesto ficticio, un jefe ha dado una conferencia y se ha llevado a su equipo de colaboradores como asistentes. Podría obtener información de cada uno de ellos para saber si realmente habían atendido a las palabras de su jefe, si habían reído o sonreído ante sus comentarios pretendidamente ingeniosos o graciosos, o si, desgraciadamente para el colaborador, se evidenciase que se había aburrido como una ostra.

Este es un muy buen ejemplo de la combinación del dato y la Inteligencia Artificial. Y de cómo nuestros gestos, nuestras sonrisas, nuestros aplausos o nuestros bostezos producto del aburrimiento son información personal, pero que si nunca nos hemos parado a pensarlo, nos pasaría totalmente desapercibido. Y sin embargo, en determinadas circunstancias, podría influir por ejemplo en una relación más o menos cordial con un compañero de trabajo.

En la actualidad hay disponible una buena panoplia de herramientas digitales que utilizan algoritmos que interpretan los gestos y las emociones implícitas en ellos:

- ***Affectiva:** desarrolla tecnología de emoción AI y análisis de afecto, utilizada en estudios de mercado, análisis de experiencia del consumidor, y desarrollo de vehículos autónomos para mejorar la seguridad al entender el estado emocional del conductor.*

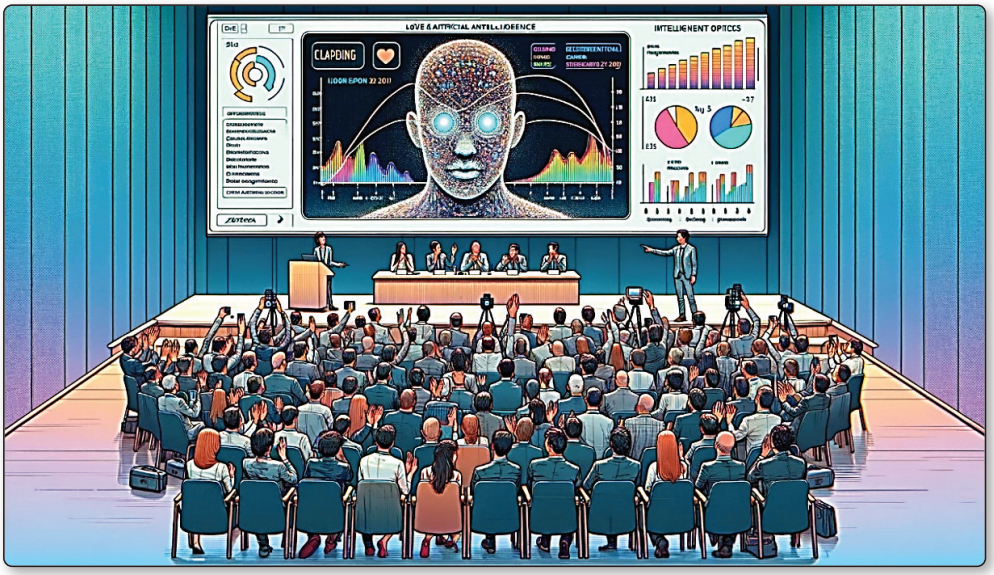
- ***Kairos:** ofrece soluciones de reconocimiento facial y de emociones destinadas a diversos sectores, incluyendo la seguridad, el marketing y la investigación de mercados.*

- ***Emotient (adquirida por Apple):** una tecnología de análisis de expresiones faciales que era capaz de detectar emociones en rostros humanos antes de ser adquirida por Apple, sugiriendo su integración en productos y servicios de la marca.*

- ***Microsoft Azure Cognitive Services:** proporciona una suite de servicios de Inteligencia Artificial, incluyendo visión por computadora y análisis de emociones, que permiten a los desarrolladores incorporar capacidades de reconocimiento emocional en sus aplicaciones.*

- ***IBM Watson Tone Analyzer:** aunque más centrado en el análisis de texto, Watson Tone Analyzer también puede ayudar a interpretar emociones y tonos en la comunicación, lo que podría complementarse con análisis visuales para una comprensión más rica de las emociones humanas.*

- ***Realeyes:*** *utiliza cámaras web para medir cómo las personas se sienten mientras visualizan contenido publicitario y de medios, proporcionando insights valiosos sobre la respuesta emocional de los espectadores.*
- ***Sightcorp DeepSight Toolkit:*** *es un software de análisis de emociones y demografía facial que proporciona datos en tiempo real sobre la edad, género, emociones, y otras métricas faciales a partir de flujos de vídeo.*



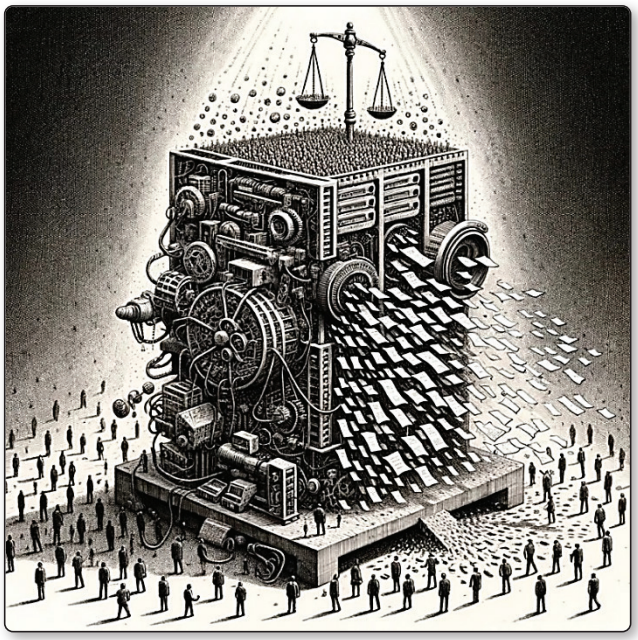
Cada vez hay más programas que se utilizan para interpretar nuestros gestos, y a través de ellos, deducir nuestras emociones. ¿Es lícito usarlos? ¿Son fiables?

Este es un ejemplo de cómo hace ya seis años, los profesionales del sector del derecho tecnológico se enfrentaban a retos sobre los riesgos a la privacidad que las novedades tecnológicas planteaban. Con el fin de ir actualizando la cobertura y seguridad legal de los ciudadanos frente a los sucesivos avances que la digitalización y el desarrollo de las redes de telecomunicación iban introduciendo en nuestras vidas, la normativa marco de Protección de Datos se fue completando con otro haz de leyes y reglamentos que a la vez que regulaban el funcionamiento de actividades o sectores concretos, y de cómo se debían usar en los mismos los datos personales. En cada una de estas legislaciones específicas se fueron incorporando como uno de los elementos más a regular el uso de algoritmos o programas que de forma automatizada recogían y procesaban datos personales. Veamos algunos ejemplos:

- Regulación de las cookies. En España a través de la Ley de Servicios de Sociedad de la Información de 2002, y sus posteriores actualizaciones.

- Reglamento europeo de 2017 sobre productos sanitarios.
- Reglamento Europeo de Ciberseguridad de 2019.
- Ley Española sobre los derechos de información a los representantes de los trabajadores sobre el uso de algoritmos en los procesos laborales. La denominada Ley de "runners", del año 2021.
- Ley Española de igualdad de trato y no discriminación en las Administraciones Públicas, de 2022.
- Reglamento Europeo de Servicios Digitales del año 2022.
- Reglamento Europeo de Gobernanza de Datos de 2022.
- Directiva Europea de Responsabilidad Civil Extracontractual de la Inteligencia Artificial, de 2022.
- Proyecto de directiva de seguridad de las máquinas, también de 2022.

En todas estas normas se aborda, -en algunos casos puntualmente, en otros casos, es el objeto completo de la norma-, la utilización de algoritmos, o como se denomina en el varias legislaciones, sistemas automatizados de toma de decisiones. Y se regula cómo debe ser su uso.



¿Es eficaz la hiper regulación legislativa?

Como vemos, la concienciación por parte de los legisladores sobre el uso correcto de la IA viene ya de lejos. Y como pueden comprobar, antes de la promulgación del Reglamento de Inteligencia Artificial, ya estaban en vigor otras numerosas normas sectoriales que regulaban el uso, cada una en su campo, de estas tecnologías.

Pero el compromiso por hacer de la Inteligencia Artificial una tecnología segura y confiable no se circunscribe a la iniciativa legislativa pública. Desde la sociedad civil, y en concreto, desde los organismos regulatorios de la calidad (ISO / IEC), llevan desde 2020 publicando una serie de normas de calidad, algunas de ellas ya certificables, que buscan procedimentar el control de riesgos y las reglas de gobernanza de estas tecnologías. Una lista ejemplificativa de estas normas sería la siguiente:

- ISO/IEC TR 24028 Tecnología de la información – Inteligencia Artificial – Visión general de la confiabilidad en Inteligencia Artificial. Mayo de 2020.
- ISO/IEC TR 24372 Tecnología de la información — Inteligencia Artificial – Visión general de los enfoques computacionales para sistemas de Inteligencia Artificial. Diciembre de 2021.
- ISO/IEC 38507 Tecnología de la información – Gobernanza de TI – Implicaciones de gobernanza del uso de la Inteligencia Artificial por las organizaciones. Abril de 2022.
- ISO/IEC 23053 Tecnología de la información – Inteligencia Artificial – Evaluación del rendimiento de clasificación de los modelos machine Learning. Junio de 2022.
- ISO/IEC 22989:2022 Tecnología de la información – Inteligencia Artificial – Conceptos y terminología de la Inteligencia Artificial. Julio de 2022.
- ISO/IEC 24668:2022 Tecnología de la información – Inteligencia Artificial – Marco de gestión de procesos para big data análisis. Noviembre de 2022.
- ISO/IEC 23894:2023 Tecnología de la información–Inteligencia Artificial – "Orientaciones sobre la gestión de riesgos". Febrero de 2023.

Como podemos ver, desde la pandemia existe un proceso acelerado de redacción de estas normas de calidad, que están íntimamente vinculadas a su vez con las legislaciones públicas están regulando la IA. Y están íntimamente vinculadas, como veremos más adelante, porque los legisladores van a exigir en un buen número de casos, que los sistemas de IA acrediten poseer certificaciones de calidad en vigor para poder operar.

1.4 INICIATIVAS EN EUROPA. RESUMEN DE LEYES VIGENTES. PRÓXIMAS LEGISLACIONES

1.4.1 Reglamento Europeo de Inteligencia Artificial

Como indicaba en el apartado anterior, éste es el mes (marzo 2024) en el que se ha aprobado definitivamente el Reglamento Europeo de Inteligencia Artificial. Como tal es un texto legal de aplicación directa en todos los países de la Unión Europea, sin tener que esperar a una transposición de sus principios legales a normas nacionales.

Sin embargo, es de esperar que, dado que un buen número de los aspectos regulados en esta norma afectan a derechos fundamentales de los ciudadanos, los países miembros hagan desarrollos legislativos complementarios, que completen los criterios de aplicación de esta norma común a cada uno de los territorios de la Unión.

Es una norma que se lleva preparando al menos desde 2019. En abril de 2021 vio la luz la propuesta de reglamento, que fue aprobada y que, tras su revisión por la cámara legislativa europea, se ha refrendado como texto legal.

¿Qué regula el Reglamento Europeo de Inteligencia Artificial (RIA)?

La comercialización, puesta en servicio y explotación (uso) de sistemas de IA.

Para lo cual establece las siguientes reglas:

- Define qué es un sistema de IA y a qué tipos de modelos o categorías de IA les afecta esta norma.
- Categoriza los sistemas IA en función a su finalidad de uso. Y en función de dicha finalidad el impacto y riesgo que su uso pudieran tener sobre los derechos y libertades de los ciudadanos y residentes en la Unión Europea. Esta clasificación es la siguiente:
  - Sistemas de IA prohibidos (Riesgo inadmisible).
  - Sistemas de IA de alto riesgo.
  - Sistemas de IA de riesgo mínimo.
  - Sistemas de IA con riesgo específico para la transparencia.
  - Sistemas de IA vinculados a los modelos de Inteligencia Artificial de uso general (GPAIS).

- Determina, en función de dicha categorización, qué sistemas de IA pueden utilizarse y cuales están inicialmente prohibidos (salvo algunas excepciones).
- Establece una serie de requisitos que deben cumplir el resto de los sistemas de IA autorizados en función del nivel de riesgo establecido en esa definición de categorías.
- Enfoca el diseño, el entrenamiento, la puesta en servicio y el uso de los sistemas de IA en función de un modelo de gestión del riesgo, que identifique los riesgos potenciales que pudiera tener el uso de cada sistema de IA, y los procedimientos preventivos destinados a minimizar al máximo cada uno de los riesgos detectados, o alternativamente procedimientos correctivos, a fin de restaurar la situación anterior a un mal uso o funcionamiento de esas IA.
- Establece qué requisitos de gestión de calidad, incluso certificación, deben cumplir los responsables del uso y/o comercialización de un sistema de IA antes de ponerlo en servicio, en función de ese nivel de riesgo y categoría legal asociada.
- Crea una autoridad de supervisión y control europea, y a su vez la posibilidad de poner en marcha en cada estado miembro de la Unión a su vez oficinas responsables de supervisión, que tienen la responsabilidad de controlar el cumplimiento y aplicación de esta norma. En este sentido es un modelo casi idéntico al establecido para la aplicación de la normativa de Protección de Datos. En España, en 2023 se promulgó la puesta en marcha de la AESIA, la Agencia Española de Supervisión de Inteligencia Artificial, que tiene su sede en A Coruña. Es precisamente a esta Agencia a la que en el futuro se podrán hacer la denuncias y reclamaciones pertinentes.
- Define una red de futuros laboratorios que colaborarán con las agencias públicas de supervisión y control en los procesos de auditar los sistemas de IA y comprobar que cumplen con la normativa. Son los denominados Organismos Comunicados.
- Establece la necesidad de registrar los sistemas de IA de alto riesgo utilizados por autoridades públicas u organizaciones que operen en su nombre.
- Crea una batería de sanciones administrativas en caso de incumplimiento de estas normas que, en los casos más graves, pueden alcanzar los 35 millones de € o el 7% de la facturación del año anterior de la empresa o entidad sancionada.

¿Cuándo entra en vigor?

La norma aprobada entra en vigor a los veinte días de su publicación en el Diario Oficial de la Unión Europea. Es decir, que ya estará en vigor en el verano o el otoño de 2024. Otra cosa es cuando es de aplicación. La norma establece una serie de tramos para la aplicación de su articulado.

- **A los seis meses** de su entrada en vigor los estados deben ir eliminando progresivamente los sistemas de IA prohibidos.
- **A los doce meses** serán aplicables a los sistemas de IA de uso general (GPAIS).
- **A los veinticuatro meses** se aplicarán a todos los sistemas de IA incluidos en la norma.
- **Solamente quedará una excepción** para algunos sistemas de alto riesgos, que se retrasa la aplicación de la normativa hasta los 36 meses de su entrada en vigor.

Como podemos ver, en un año programas tan conocidos como ChatGPT tendrán que haberse adecuado a esta norma. ¿Es así? ¿Pero ChatGPT no es un programa de una compañía norteamericana? ¿Por qué debe cumplir con esta regulación? Muy sencillo, por los **efectos de extraterritorialidad de la norma europea**. La norma no tiene en cuenta desde qué país se presta el servicio o se usa el sistema de IA. Lo que tiene en cuenta es si se usa o afecta a un ciudadano que resida en la UE o se ofrece dentro del territorio de la UE.

Más adelante, en el apartado "ANEXO AL CAPÍTULO I", analizaremos con detalle esta norma.

1.4.2 Reglamento Europeo de Protección de Datos

Como vimos anteriormente, el legislador europeo, ya había tomado conciencia de la importancia de informar al interesado[1] del uso de sus datos personales para realizar tratamientos automatizados de dichos datos que pudieran afectar a sus derechos. Y esa regulación se plasmó en la ya derogada Directiva de 1995.

---

1 Se denomina con el término "interesado" a cualquiera de nosotros que, como particulares, aceptamos ceder nuestros datos personales.

Cuando se abordó la redacción del Reglamento Europeo de Protección de Datos, actualmente en vigor; lógicamente, el legislador no se olvidó de este derecho de los ciudadanos; ahora con mayor justificación dado la extensión y el uso del Big Data y la elaboración de perfiles que se ha hecho extensiva a toda la economía digital.

Por ello, el artículo 13, apartado 2 párrafo f) dice lo siguiente:

"Información que deberá facilitarse cuando los datos personales se obtengan del interesado.

2. Además de la información mencionada en el apartado 1 (los datos básicos de los fines de uso de los datos y quienes son los responsables en el tratamiento de esos datos), el responsable del tratamiento facilitará al interesado, en el momento en que se obtengan los datos personales, la siguiente información necesaria para garantizar un tratamiento de datos leal y transparente:

f) La existencia de decisiones automatizas, incluida la elaboración de perfiles, a que se refiere el artículo 22, apartados 1 y 4, y, al menos en tales casos, información significativa sobre la lógica aplicada, así como la importancia y las consecuencias previstas de dicho tratamiento para el interesado."

Es decir, que si la empresa, persona o entidad que solicita nuestros datos personales piensa hacer tratamientos automatizados de los mismos, usando algoritmos que procesan dichos datos y tomen u ofrezcan decisiones sobre dicho tratamiento, **deberá informarnos previamente** de ese hecho. Y sobre todo **deberá informarnos del impacto que dicho tratamiento puede tener sobre nosotros.**

Es una de las consecuencias del deber de transparencia al que está obligado el responsable de tratamiento y que, como veremos en el apartado para lectores avanzados, también se traslada al Reglamento de Inteligencia Artificial.

Este mismo deber se traslada a los responsables del tratamiento que hayan obtenido los datos personales no directamente del interesado, sino a través de terceros que se los hayan cedido. Aparece regulado en el artículo 14.2.g) del Reglamento General de Protección de Datos.

Ambos artículos se refieren a su vez al artículo **22 del Reglamento General de Protección de Datos,** que dice así:

"Decisiones individuales automatizadas, incluida la elaboración de perfiles

1. Todo interesado tendrá derecho a no ser objeto de una decisión basada únicamente en el tratamiento automatizado, incluida la elaboración de perfiles, que produzca efectos jurídicos en él o le afecte significativamente de modo similar."

Una vez planteado el principio general, el artículo en su apartado 2 establece las excepciones a la aplicación de este principio:

> *2. El apartado 1 no se aplicará si la decisión:*
>
> *a) es necesaria para la celebración o la ejecución de un contrato entre el interesado y un responsable del tratamiento;*
>
> *b) está autorizada por el Derecho de la Unión o de los Estados miembros que se aplique al responsable del tratamiento y que establezca asimismo medidas adecuadas para salvaguardar los derechos y libertades y los intereses legítimos del interesado, o*
>
> *c) se basa en el consentimiento explícito del interesado.*

Como podemos ver, al final, en muchas ocasiones de nuestras vidas, vamos a tener que aceptar ese tratamiento automatizado si queremos acceder a muchos servicios, y por lo tanto, tendremos que aceptar explícitamente dicho tratamiento automatizado.

En el apartado tercero de este artículo 22 deja abierta la puerta a solicitar la intervención humana en ese proceso de valoración automática o a impugnarlo, **manteniendo el derecho de impugnación de valoraciones** que, como vimos al principio de este capítulo, estaba consagrado en nuestra Ley Orgánica de protección de datos de 1999.

> *3. En los casos a que se refiere el apartado 2, letras a) y c), el responsable del tratamiento adoptará las medidas adecuadas para salvaguardar los derechos y libertades y los intereses legítimos del interesado,* como mínimo el derecho a obtener intervención humana *por parte del responsable, a expresar su punto de vista y* a impugnar la decisión.

El apartado 4 de este artículo 22 pasa habitualmente desapercibido en los análisis legales, pero a día de hoy, a la luz de la evolución de las tecnologías inteligentes de reconocimiento por imagen y biométrico, cobra especial importancia:

> *1. Las decisiones a que se refiere el apartado 2* no se basarán en las categorías especiales de datos personales contempladas en el artículo 9*, apartado 1, salvo que se aplique el artículo 9, apartado 2, letra a) o g), y se hayan tomado medidas adecuadas para salvaguardar los derechos y libertades y los intereses legítimos del interesado.*

¿Y cuáles son esos datos especialmente protegidos que enumera el **artículo 9 del Reglamento general de Protección de Datos?**

- Origen étnico o racial.
- Las opiniones políticas.
- Las convicciones religiosas o filosóficas.
- La afiliación sindical.
- El tratamiento de datos genéticos.
- Los datos biométricos dirigidos a identificar de manera unívoca a una persona física.
- Datos relativos a la salud.
- Datos relativos a la vida o a la orientación sexual de una persona física.

Pues bien, la Ley prohíbe expresamente que se puedan realizar tratamientos automatizados de datos usando estas categorías; y por lo tanto elaborando perfiles digitales de los ciudadanos en base a estos datos especialmente protegidos.

Salvo….

Que el propio interesado haya dado su consentimiento.

Que el tratamiento sea necesario por el interés público esencial, término suficientemente amplio que podría dar pie a futuros abusos.

1.4.3 Reglamento Europeo de Servicios Digitales

El Reglamento de Servicios Digitales (también conocido como «Ley de Servicios Digitales») es una norma que obliga a las empresas de servicios digitales de toda la UE a rendir cuentas por los contenidos publicados en sus plataformas.

Promulgado el pasado 19 de octubre de 2022. Está plenamente en vigor desde el 17 de febrero de 2024.

El Reglamento de Servicios Digitales busca ofrecer entornos de relación y comercio online más seguros y respetuosos con los derechos de las empresas y los ciudadanos.

Afecta principalmente a las grandes plataformas digitales que comercializan, distribuyen contenidos productos y servicios, aunque también a los grandes buscadores online. Grandes tiendas en línea, plataformas de viajes y alojamientos en línea, plataformas de contenidos, plataformas de servicios profesionales, buscadores online. En definitiva, y en palabras de la propia UE busca:

- *"garantizar que los usuarios de servicios digitales tengan acceso a productos seguros y proteger los derechos fundamentales de los usuarios;*
- *permitir una competencia libre y leal en los sectores digitales para impulsar la innovación y el crecimiento."*

Y evitar problemas de discriminación profesional o comercial o la extensión de noticias falsas a través de Internet, entre otros objetivos.

En su artículo 14 apartado 1, esta norma exige a las plataformas de contenidos digitales, venta online, redes sociales o buscadores un deber de transparencia asociado al uso de algoritmos en sus procesos de relación con los clientes o usuarios:

1. Los prestadores de servicios intermediarios incluirán en sus condiciones generales información sobre cualquier restricción que impongan en relación con el uso de su servicio respecto de la información proporcionada por los destinatarios del servicio. *Esta información deberá incluir datos sobre cualesquiera políticas, procedimientos, medidas y herramientas empleadas para moderar los contenidos,* incluidas la toma de decisiones mediante algoritmos y la revisión humana, *así como sobre las normas de procedimiento de su sistema interno de gestión de reclamaciones. Se expondrá en lenguaje claro, sencillo, inteligible, accesible al usuario e inequívoco, y se hará pública en un formato fácilmente accesible y legible por máquina.*

Es decir, que en el momento de informarnos cuando vamos a acceder a uno de estos servicios, deberán indicar si usan determinados tipos de algoritmos que puedan limitar la información que se nos ofrece, y en qué medida afecta a dicha información u oferta presentada al usuario.

Hay que tener presente que a día de hoy, las grandes corporaciones propietarias de los buscadores ya han sido sancionadas por alterar los algoritmos de búsqueda a fin de favorecer a determinadas empresas (o informaciones) o perjudicar o limitar el acceso de otros operadores económicos. O el acceso a otras fuentes de información o de noticias.

Igualmente, como veremos más adelante en el capítulo 5, las grandes plataformas de búsqueda de empleo y redes sociales destinadas a contactos profesionales, utilizan algoritmos que pueden condicionar nuestras posibilidades de encontrar las ofertas de empleo que deseamos, en función a la creación de un perfilado usando datos personales nuestros, sin nuestro conocimiento y consentimiento previo, que pueden presentar ante nuestros ojos determinadas ofertas u oportunidades de empleo, y descartar otras en función de la decisión del algoritmo, sin que seamos consciente de ello.

ANEXO PARA LECTORES AVANZADOS: los filtros burbuja

Esos peligrosos algoritmos que limitan el acceso a la información y pueden condicionar nuestra opinión

Uno de los grandes riesgos al que nos enfrentamos es por el uso de algoritmos que, creando un perfil de nuestros gustos, aficiones, conocimientos, renta, entorno social o incluso ideología, nos presenta noticias, ofertas o contenidos que pueden limitar el acceso a información, o a ofertas comerciales, y que en el peor de los casos, pueden ir radicalizando la opinión pública, al trasladar a través de los motores de búsqueda o grandes plataformas de información en la que solamente se da una visión de la realidad, imposibilitando que el lector compare la interpretación de una noticia con varias versiones editoriales, por ejemplo.

A este hecho se le denomina Erosión Epistémica[2], y es uno de los grandes problemas a los que se enfrenta nuestras sociedades digitales en estos momentos.

Los filtros burbuja, que empezaron a utilizarse a partir de 2009 para lograr una experiencia más personalizada del usuario en los motores de búsqueda, han derivado en un arma de doble filo en donde se generan entornos cerrados, endogámicos, donde solo hay una visión, solo hay una opinión o solamente se ofrecen alternativas muy limitadas al consumo o a la búsqueda de contenidos, información o soluciones.

2 Sorprende cómo cien años más tarde, tesis como las de **Walter Benjamín y Theodor W. Adorno**, **sobre la reformulación de la verdad están plenamente vigentes**. La catástrofe que supone en nuestros días de la construcción del pensamiento y la opinión pública, contaminada de noticias falsas y una avalancha de juicios de valor sin sentido, anida en la **subversión del concepto de verdad**, al contraponer el concepto de **verdad-racionalidad**, con los conceptos de **verdad-negativa** y **verdad y dolor**, frente a los que el ciudadano moderno huye, y algoritmos como los filtros burbuja ayudan al individuo moderno en esa huida de la realidad dolorosa.

Los filtros burbuja.

1.4.4 Reglamento Europeo relativo a las máquinas

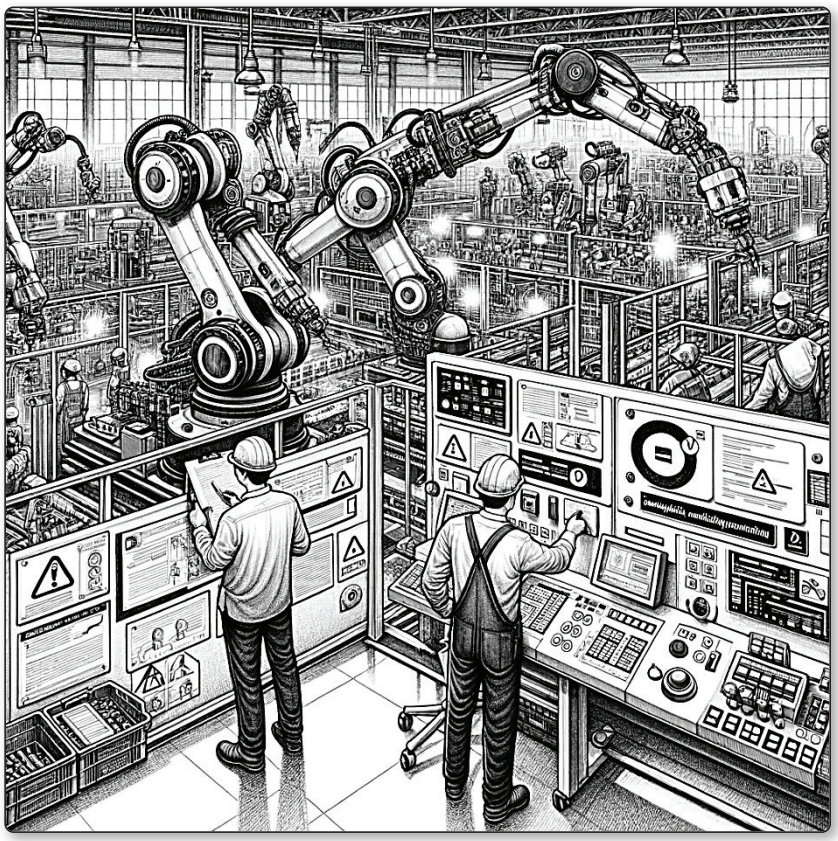
Este reglamento está aún en fase de propuesta y no ha entrado en vigor. Busca garantizar la calidad de las máquinas que se distribuyen en el mercado único y su seguridad. En la búsqueda de este fin aborda en profundidad los aspectos de seguridad en el funcionamiento de la maquinaría, la robótica, el diseño y puesta en servicio de productos y el Internet de las cosas y los productos conectados, y su relación directa con el uso de sistemas de IA asociados a estos productos.

El legislador es consciente de la existencia de un Reglamento Europeo de Inteligencia Artificial, y como consecuencia, la regulación que esta norma hace de la utilización de la IA en las máquinas es siempre subsidiaria a las normas generales que establece el Reglamento de IA, y busca cubrir determinados aspectos específicos que el Reglamento de IA no puede prever por su carácter general.

Artículo 9

Cuando las máquinas y sus partes y accesorios contengan un sistema de Inteligencia Artificial al que se apliquen los requisitos esenciales de salud y seguridad del Reglamento (UE) (de Inteligencia Artificial), el presente Reglamento solo se aplicará, en relación con dicho sistema de Inteligencia Artificial, ***en lo que respecta a su integración segura en la máquina completa****, a fin de no comprometer la seguridad del producto en su conjunto.*

En definitiva, que los fabricantes de maquinaria que integren en la misma, para su funcionamiento, sistemas de IA, deberán hacer pasar a dichos sistemas por un control de calidad que garantice la seguridad de la máquina. Y en muchos casos ese control de calidad quedará definido por el Reglamento Europeo de IA.



El porcentaje de inversiones en desarrollo de sistemas de IA que se aplican a la robótica, maquinaria e instalaciones industriales, la denominada Inteligencia Artificial Situada, es superior al destinado a programas comerciales, plataformas digitales, software de contenidos o sistemas de propósito general (Inteligencia Artificial No Situada), pero la visibilidad social de estos sistemas de IA situada es casi nula.

1.4.5 Directiva Europea de Responsabilidad Civil Extracontractual de la Inteligencia Artificial

Con esta norma terminamos de ofrecer una primera visión de las iniciativas asociadas a dar cobertura legal y protección de los derechos de los ciudadanos por los efectos del uso de la Inteligencia Artificial.

A diferencia de las normas anteriores, esta normativa no es de aplicación directa. Como directiva lo que establece es un marco regulatorio general que deberá trasponerse a la legislación de cada uno de los países de la Unión Europea.

Esta directiva, que igualmente ha tenido muy poca visibilidad a nivel general, trata sin embargo un tema crucial en la defensa de los derechos e intereses de los ciudadanos frente a los efectos de un mal funcionamiento de un sistema de IA, y como consecuencia, las responsabilidades a las que la organización o empresa que use ese sistema de IA defectuoso, o eventualmente su diseñador o fabricante, deba hacer frente.

El problema de muchos sistemas de IA es que es muy difícil demostrar por parte de una persona afectada que el mismo ha funcionado de forma defectuosa. Es a causa del denominado efecto de la "caja negra". Igualmente es difícil seguir el rastro de los auténticos responsables en el uso de un sistema de IA.

ANEXO PARA LECTORES AVANZADOS: ¿Qué es una "caja negra" en la Inteligencia Artificial?

Es un modelo de IA, es decir, el conjunto de algoritmos agrupados que constituyen un sistema determinado de IA que está en funcionamiento, que ofrecen un resultado o toman una decisión sin explicar cómo han llegado hasta esa conclusión. O al menos mostrar cual ha sido el proceso de toma de decisión. Los procesos internos de análisis de datos y toma de decisiones usados por esos algoritmos y los factores que han manejado para ponderar dichos datos quedan totalmente ocultos.

En otras palabras: hay una total falta de transparencia a la hora de analizar el funcionamiento de esta tecnología. Ni siquiera los programadores que diseñaron el modelo de IA, ni sus administradores son capaces de conocer cómo y por qué ha tomado una decisión y no otra.

Caso de ejemplo: un modelo de caja negra que puntúe el riesgo financiero de los clientes funciona en base a unos datos que se ingresan sobre ese cliente, que pueden abarcar desde datos personales como la edad, su situación laboral, si tiene familia dependiente, si está casada esa persona o no; pasando por lo datos más objetivos como el nivel de ingresos, si tiene propiedades, si tiene otros préstamos