INFORMÁTICA Y DERECHO

REVISTA IBEROAMERICANA DE DERECHO INFORMÁTICO
(SEGUNDA ÉPOCA)

FEDERACIÓN IBEROAMERICANA DE ASOCIACIONES
DE DERECHO E INFORMÁTICA

ISSN 2530-4496 – N.º 15 VOL. 1, 2024, PP. 99-107

LAS OBLIGACIONES DE TRANSPARENCIA PARA ELIMINAR LOS RIESGOS DE LOS SISTEMAS DE INTELIGENCIA ARTIFICIAL

TRANSPARENCY OBLIGATIONS TO ELIMINATE THE RISKS OF ARTIFICIAL INTELLIGENCE SYSTEMS

Elisa Palomino Angeles

Universidad Autónoma Metropolitana

Resumen

Con la evolución de los sistemas de inteligencia artificial, se ha masificado una diversidad de riesgos, por ello, el Parlamento Europeo crea un Reglamento Europeo sobre Inteligencia artificial, en el cual se establecen normas armonizadas en la materia, en la cual, de acuerdo al enfoque del riesgo, se clasifican los riesgos en inaceptables, de alto riesgo, riesgo limitado y riesgo bajo o mínimo. Asimismo, se contemplan los riesgos genéricos y sistémicos, y se regulan las obligaciones de transparencia, las cuales analizamos como nuestro objeto de investigación para poder determinar si existen lagunas, incoherencias o problemas semánticos que hagan a la norma ineficaz para eliminar los riesgos que se originan con el uso de la inteligencia artificial. Entre esas obligaciones encontramos que el proveedor deberá informar al usuario que está interactuando con un sistema de IA, excepto cuando sea evidente que lo está haciendo con una persona física, razonablemente informada, atenta y perspicaz teniendo en cuenta las circunstancias y contexto.

Palabras clave

Obligaciones, transparencia, riesgos, inteligencia artificial.

Abstract

With the evolution of artificial intelligence systems, a diversity of risks has become widespread, therefore, the European Parliament created a European Regulation on Artificial Intelligence, which establishes harmonized rules on artificial intelligence, in which risks are classified as unacceptable, high risk, limited risk, and low or minimal risk. Likewise, generic and systemic risks are contemplated, and transparency obligations are regulated, which we analyze as our object of research in order to determine if there are gaps, inconsistencies or semantic problems that may cause the regulation to be ineffective in eliminating the risks that arise from the use of artificial intelligence. Among these obligations, we find that the provider must inform the user that he or she is interacting with an AI system, except when it is evident that you are doing so with a natural person, reasonably informed, attentive and discerning taking into account circumstances and context.

Keywords

Obligations, transparency, risks, artificial intelligence.

Introducción

En la actualidad, la complejidad digital de los sistemas provoca que se presenten más dificultades para desarrollar sistemas jurídicos que nos permitan contar con el mínimo de protección de los usuarios o consumidores de los sistemas de inteligencia artificial. Hoy se requieren normas eficaces en las cuales se pueda establecer instrumentos jurídicos que materialicen las obligaciones de la transparencia que les corresponde a los proveedores de sistemas de inteligencia artificial, como un derecho para eliminar algunos de sus riesgos. ¿Cuáles son las obligaciones que evitarán o eliminarán los riesgos de uso y aplicación de los sistemas de inteligencia artificial? ¿Se podrá establecer la igualdad o equidad en las relaciones asimétricas que hoy se presentan entre los usuarios y los proveedores de estos sistemas? Es la capacidad económica de los sujetos en esta relación asimétrica uno de los elementos que favorece la desigualdad entre los mismos proveedores en relaciones desiguales.

La inteligencia artificial

El sistema de inteligencia es una complejidad digital en la que interactúa una diversidad de elementos, tecnologías digitales, sujetos, instituciones, gobiernos, la entropía que producen estas relaciones jurídicas, tanto simétricas como asimétricas, la cual ha originado una necesidad de establecer modelos de seguridad para los sistemas de inteligencia artificial, en los cuales las normas sean eficaces de tal manera que no se vulneren derechos fundamentales.

En ese orden de ideas tenemos que iniciar primero con los aspectos fundamentales, como son algunas definiciones del Reglamento Europeo de Inteligencia Artificial (Ley sobre Inteligencia Artificial), promulgado el 13 de marzo de 2024 por el Parlamento Europeo, el cual establece la definición legal de inteligencia artificial de la siguiente manera:

> A los efectos del presente Reglamento, se entenderá por: 1) «sistema de IA»: un sistema basado en una máquina diseñado para funcionar con distintos niveles de autonomía, 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 información de salida, como predicciones, contenidos, recomendaciones o decisiones, que puede influir en entornos físicos o virtuales. (Art. 3.1).

Uno de los problemas ha sido definir qué es un sistema de inteligencia artificial, porque existe una multiplicidad de significados desde diversas perspectivas: jurídica, económica, y tecnológica, y es en la Ley sobre Inteligencia Artificial del Parlamento Europeo que se establece la referida definición legal, la cual fue cuestionada por la Big Data Value Association,

> una organización internacional sin ánimo de lucro impulsada por la industria. Subrayó que la definición de sistemas de IA era bastante amplia y abarcaría mucho más de lo que se entiende subjetivamente como IA, incluidos los algoritmos de búsqueda, clasificación y enrutamiento más simples, que en consecuencia estarían sujetos a nuevas reglas. (Parlamento Europeo, 2024).

De lo que se desprende que se incluye más sistemas de inteligencia artificial, e incluso las de más bajo nivel, contemplados en la referida ley, pero también encontramos la el argumento de AmCham, la Cámara de Comercio Americana en la UE, que sugirió

> evitar el exceso de regulación mediante la adopción de una definición más estrecha de los sistemas de IA, centrándose estrictamente en las aplicaciones de IA de alto riesgo (y no extendida a las aplicaciones de IA que no son de alto riesgo, o al *software* en general). (Parlamento Europeo, 2024).

Por lo que se puede desprender que puede ser muy amplia la definición de inteligencia artificial, y en este caso se pretende evitar un exceso de regulación a sistemas de inteligencia con mínimo riesgo.

Es interesante lo que refiere AmCham porque los expertos en los sistemas inteligentes pueden realizar una clasificación de estos, pero hay que considerar que esa multiplicidad de sistemas no solo se debe clasificar de acuerdo al riesgo, sino que también se debe considerar más categorías, como la potencia del sistema y el nivel de autonomía para alcanzar metas, la capacidad de aprender, de adaptarse, es decir, aquellas que tomen decisiones sin control y con alto riesgo, porque no todo es inteligencia artificial.

Tipos de riesgos

En el referido Reglamento Europeo se regula las definiciones de riesgo, que se entiende como la combinación de la probabilidad de que se produzca un perjuicio y la gravedad de dicho perjuicio (art. 3. inciso 2), se tiene que determinar el daño causado y la gravedad de este para poder determinar su pago de indemnización, que es considerado como genérico, ahora entendemos como «perjuicio la privación de cualquier ganancia lícita, que debiera haberse obtenido con el cumplimiento de la obligación» (México, Suprema Corte de Justicia, 1967). Es decir, una lesión al patrimonio de los usuarios o consumidores de sistemas de inteligencia artificial, donde la responsabilidad, tanto civil como penal, deberá reparar el perjuicio; y, por otra parte, el término gravedad como la importancia que tiene el perjuicio.

Ahora bien, también se regula el «riesgo sistémico»:

> un riesgo específico de las capacidades de gran impacto de los modelos de IA de uso general, que tienen unas repercusiones considerables en el mercado de la Unión debido a su alcance o a los efectos negativos reales o razonablemente previsibles en la salud pública, la seguridad, la seguridad pública, los derechos fundamentales o la sociedad en su conjunto, que puede propagarse a gran escala a lo largo de toda la cadena de valor.

En este orden de ideas, el Parlamento Europeo, con base en el riesgo, regula la clasificación de estos, entre los cuales incluye: a) riesgos inaceptables o prohibidos, b) sistemas de inteligencia artificial con alto riesgo, c) sistemas de inteligencia artificial con especificaciones, y d) sistemas de inteligencia artificial con mínimo riesgo o sin riesgo. De ellos, solo mencionaremos a los dos primeros por su trascendencia.

a) Riesgo inaceptable:

> (…) los sistemas de IA de riesgo inaceptable son los que se consideran una amenaza para las personas y serán prohibidos. Incluyen:
>
> Manipulación cognitiva del comportamiento de personas o grupos vulnerables específicos: por ejemplo, juguetes activados por voz que fomentan comportamientos peligrosos en los niños
>
> Puntuación social: clasificación de personas en función de su comportamiento, estatus socioeconómico o características personales Sistemas de identificación biométrica en tiempo real y a distancia, como el reconocimiento facial (…). (Parlamento Europeo, 2023).

Además, en relación con el art. 1, el art. 5 regula las «Prácticas de IA prohibidas»:

> 1. Quedan prohibidas las siguientes prácticas de IA: a) La introducción en el mercado, la puesta en servicio o la utilización de un sistema de IA que se sirva de técnicas subliminales que trasciendan la conciencia de una persona o de técnicas deliberadamente manipuladoras o engañosas con el objetivo o el efecto de alterar de manera sustancial el comportamiento de una persona o un grupo de personas, mermando de manera apreciable su capacidad para tomar una decisión informada y haciendo que una persona tome una decisión que de otro modo no habría tomado, de un modo que provoque, o sea probable que provoque, perjuicios considerables a esa persona, a otra persona o a un grupo de personas (…). (Parlamento Europeo, 2024).

Es un acierto el haber prohibido la manipulación cognitiva y salvaguardado la libertad de expresión, el consentimiento informado y la salud mental del individuo. Resulta acertado y necesario establecer excepciones en el caso de los sistemas de identificación biométrica en tiempo real y a distancia, sobre todo tratándose de casos de seguridad pública.

b) Sistemas de inteligencia artificial con alto riesgo: son aquellos permitidos al sujeto siempre que cumpla los requisitos del reglamento de IA y la evaluación de conformidad (Ministerio de Asuntos Económicos y Transformación Digital, s. f.).

c) Sistemas de inteligencia artificial con especificaciones: son aquellos permitidos pero, tanto proveedores como usuarios, están sujetos a cumplir con la transparencia.

d) Sistemas de inteligencia artificial con mínimo riesgo o sin riesgo: son aquellos permitidos.

Las obligaciones de transparencia como elementos para eliminar los riesgos en los sistemas de IA

El derecho a la información es un derecho fundamental que genera obligaciones de transparencia por parte de proveedores y usuarios, pero ¿qué entendemos por transparencia? De acuerdo con la OCDE, la transparencia es un concepto relacionado con la posibilidad de que la información real de una empresa, gobierno u organización sea consultada por los diferentes sujetos afectados por

ella, de tal modo que estos puedan tomar decisiones con conocimiento de causa y sin asimetría de información (Perramon, 2013).

En el Título IV, artículo 50, «Obligaciones de transparencia de los proveedores y usuarios de determinados sistemas de IA», se señala:

> 1. Los proveedores garantizarán que los sistemas de IA destinados a interactuar directamente con personas físicas se diseñen y desarrollen de forma que las personas físicas de que se trate estén informadas de que están interactuando con un sistema de IA, excepto cuando resulte evidente desde el punto de vista de una persona física razonablemente informada, atenta y perspicaz, teniendo en cuenta las circunstancias y el contexto de utilización. Esta obligación no se aplicará a los sistemas de IA autorizados por ley para detectar, prevenir, investigar o enjuiciar infracciones penales, con sujeción a las garantías adecuadas para los derechos y libertades de terceros, salvo que estos sistemas estén a disposición del público para denunciar una infracción penal.

Es importante resaltar que este artículo se refiere a la interacción que tiene las personas físicas con un sistema de inteligencia artificial, en la cual la transparencia consiste en que el proveedor deberá informar que se está interactuando con un sistema de IA, a menos de que la interacción sea con un individuo razonablemente informado, atento y perspicaz, teniendo en cuenta su entorno y particularidades.

Consideremos que son ambiguas las palabras que se utilizaron como candados: tener una observación aguda y penetrativa, razonablemente informada, que una persona esté acorde a su edad y escolaridad, una visión aguda y penetrativa o extremadamente observadora (Real Academia Española, s.f.).

¿Qué parámetros se necesita para determinar la obligación de transparencia que deben tener los proveedores? En las relaciones asimétricas, como es el caso que nos ocupa, entre usuarios y proveedores, se requiere darles mayores derechos a los usuarios, por las desventajas económicas y de conocimiento que existen entre ellos. Aunado a que se requiere de conocimiento tecnológico muy especializado para poder ser razonablemente informado, esto conlleva el conocimiento de expertos, por ejemplo: no se podrá detectar que se está hablando con una máquina de IA y no con un humano, porque se pueden reproducir las voces humanas sin que la persona se percate de la situación.

Continuando con el análisis del número 2 del art. 50 referido, tenemos que dice:

> 2. Los proveedores de sistemas de IA, entre los que se incluyen los sistemas de IA de uso general, que generen contenido sintético de audio, imagen, vídeo o texto, velarán por que la información de salida del sistema de IA esté marcada en un formato legible por máquina y que sea posible detectar que ha sido generada o manipulada de manera artificial. Los proveedores velarán por que sus soluciones técnicas sean eficaces, interoperables, sólidas y fiables en la medida en que sea técnicamente viable, teniendo en cuenta las particularidades y limitaciones de los diversos tipos de contenido, los costes de aplicación y el estado actual de la técnica generalmente reconocido, según se refleje en las normas técnicas pertinentes. Esta obligación no se aplicará en la medida en que los sistemas de IA desempeñen una función de apoyo a la edición estándar

> o no alteren sustancialmente los datos de entrada facilitados por el responsable del despliegue o su semántica, o cuando estén autorizados por ley para detectar, prevenir, investigar o enjuiciar infracciones penales. (Parlamento Europeo, 2024).

Este se refiere a la obligación de transparencia que tiene el proveedor de los sistemas de inteligencia artificial de uso general que generan contenido sintético de audio, imagen, video o texto. Dice que «velará», consideramos que no es la palabra adecuada, más bien es «vigilará», aunque no se utiliza esta palabra, sino que velará por que la información del sistema de IA de salida esté marcada por el proveedor en formato legible y detectable sobre la generación o manipulación artificial, esto es un gran acierto, pero resulta ser que, como en el punto anterior, se establece una diversidad de excepciones que no tienen un lenguaje sencillo para el usuario, sino para los expertos en sistemas de inteligencia artificial, y nuevamente se utiliza excepciones a las reglas generales, lo cual podrá dar lugar a evadir responsabilidades. En las excepciones están la edición estándar o modificación sustancial de los datos de entrada.

«El formato legible por máquina y que sea posible detectar que ha sido generada o manipulada de manera artificial»: al proveedor, en esta frase, se le da la posibilidad, mas no la obligación, de tener un formato que detecte que la información ha sido generada o manipulada de manera digital, es deber tener una solución técnica viable, eficaz interoperable, sólida. Consideramos que estos requisitos deberán estar establecidos dentro de las normas técnicas pertinentes, los estándares internacionales no deberán estar sujetos a la posibilidad del proveedor, si realmente se quiere prevenir el riesgo informático, ya que si bien es cierto que existen diversas empresas de distintos niveles económicos y tecnológicos, también es cierto que será necesario tener un fondo económico para apoyar o financiar a pequeñas empresas para que puedan cumplir con estas obligaciones de transparencias.

Observamos que al establecer en la medida que sea «técnicamente viable, teniendo en cuenta las particularidades y limitaciones de los diversos tipos de contenido, los costes de aplicación y el estado actual», se ayudará los proveedores de pymes, lo cual fomentará la inalterabilidad de los datos de entrada facilitados por el responsable.

> 3. Los responsables del despliegue de un sistema de reconocimiento de emociones o de un sistema de categorización biométrica informarán del funcionamiento del sistema a las personas físicas expuestas a él y tratarán sus datos personales de conformidad con los Reglamentos (UE) 2016/679 y (UE) 2018/1725 y con la Directiva (UE) 2016/680, según corresponda. Esta obligación no se aplicará a los sistemas de IA utilizados para la categorización biométrica y el reconocimiento de emociones que hayan sido autorizados por ley para detectar, prevenir e investigar infracciones penales, con sujeción a las garantías adecuadas para los derechos y libertades de terceros y de conformidad con el Derecho de la Unión. (Parlamento Europeo, 2024).

En este punto tres, en los sistemas de reconocimientos de emociones o categorización biométrica, el responsable deberá informar del despliegue sobre su funcionamiento a las personas expuestas.

> 4. Los responsables del despliegue de un sistema de IA que genere o manipule imágenes o contenidos de audio o vídeo que constituyan una ultrafalsificación harán público que estos contenidos o imágenes han sido generados o manipulados de manera artificial (…). Los responsables del despliegue de un sistema de IA que genere o manipule texto que se publique con el fin de informar al público sobre asuntos de interés público divulgarán que el texto se ha generado o manipulado de manera artificial. Esta obligación no se aplicará cuando el uso esté autorizado por ley para detectar, prevenir, investigar o enjuiciar infracciones penales, o cuando el contenido generado por IA haya sido sometido a un proceso de revisión humana o de control editorial y cuando una persona física o jurídica tenga la responsabilidad editorial por la publicación del contenido. (Parlamento Europeo, 2024).

Generación o manipulación de imágenes, audio, video o *deepfakes*: información pública sobre su origen artificial. Si se trata de informar al público sobre asuntos de interés público, el responsable del despliegue debe informar sobre su origen salvo que haya supervisión humana o control editorial y una persona tenga responsabilidad editorial por la publicación.

Información suministrada de manera clara y distinguible en las primeras interacciones y exposiciones: es muy acertado el regular esta obligación de transparencia a la inteligencia generativa, debido a las conductas ilícitas que se pueden generar con el mal uso de esta.

Conclusiones

La transparencia específica sólo aplica a los sistemas de inteligencia artificial generativa, consideramos que deberá revisarse la clasificación de los sistemas de IA para proteger a los usuarios, que son los más vulnerables en esta cadena de valores. Además, si bien es cierto que establece obligaciones tanto para los proveedores como para los responsables del despliegue, las cuales consideramos que son pertinentes, también es cierto que establece muchas excepciones a las obligaciones, por las cuales se podrá evadir la responsabilidad civil o penal, debido a que algunas de estas normas son insuficientes, vagas y ambiguas. Las obligaciones de transparencia pueden ser eficaces si se perfeccionan las normas, sobre todo las excepciones a la regla general.

Referencias bibliográficas

México. Suprema Corte de Justicia. (1967). *Diferencia entre daño y perjuicio*. https://sjf2.scjn.gob.mx/detalle/tesis/258965

Ministerio de Asuntos Económicos y Transformación Digital. (s. f.). *El Reglamento Europeo de IA, en resumen*.

https://portal.mineco.gob.es/es-es/digitalizacionIA/sandbox-IA/Documents/20220919_Resumen_detallado_Reglamento_IA.pdf

https://www.europarl.europa.eu/topics/es/article/20230601STO93804/ley-de-ia-de-la-ue-primera-normativa-sobre-inteligencia-artificial

https://www.europarl.europa.eu/thinktank/es/document/EPRS_BRI(2021)698792

https://accid.org/wp-content/uploads/2018/10/La_transparencia._Concepto_evolucion_y_retos_a.pdf

Parlamento Europeo. (2023). *Ley de IA de la UE: primera normativa sobre inteligencia artificial*. https://www.europarl.europa.eu/topics/es/article/20230601STO93804/ley-de-ia-de-la-ue-primera-normativa-sobre-inteligencia-artificial

Parlamento Europeo. (2024). *Artificial Intelligence Act*. https://www.europarl.europa.eu/thinktank/es/document/EPRS_BRI(2021)698792

Perramon, J. (2013). La transparencia: concepto, evolución y retos actuales. *Revista de Contabilidad y Dirección*, *16,* 11-27. https://accid.org/wp-content/uploads/2018/10/La_transparencia._Concepto_evolucion_y_retos_a.pdf

Real Academia Española. (s. f.). Perspicacia. *Diccionario de la lengua española*. https://dle.rae.es/perspicacia