

La inteligencia artificial generativa y la justificación en la decisión judicial

Elisa Palomino Angeles*
Óscar Herrera Alcántara**
Amir García Villalpando***

Resumen:

La decisión judicial hoy en día enfrenta un nuevo desafío, la utilización de la inteligencia artificial generativa como creadora de hechos o actos jurídicos irreales, o reales con modificaciones o alteraciones, transformó la realidad a una artificial o sintética. Los medios de prueba han aumentado y el análisis del elemento tecnológico es indispensable para justificar la acreditación de un hecho o acto jurídico a través de las pruebas que son valoradas por los juzgadores. Estos medios de prueba requieren ser debidamente valorados, por lo que pretendemos determinar la necesidad de establecer los criterios mínimos de certidumbre, fiabilidad y seguridad jurídica que deberá considerar un juzgador para justificar su decisión judicial, buscando lo válido, verdadero y justo en el elemento científico-tecnológico.

Abstract:

The judicial decision today faces a new challenge, the use of generative artificial intelligence as a creator of unreal facts or legal acts, or real with modifications or alterations, transformed reality to an artificial or synthetic one. The means of proof have increased and the analysis of the technological element is essential to justify the accreditation of a fact or legal act through the evidence that is evaluated by the judges. These means of proof need to be duly assessed, so we intend to determine the need to establish the minimum criteria of certainty, reliability and legal certainty that a judge must consider to justify his judicial decision, looking for what is valid, true and fair in the scientific technological element.

* Profesora-Investigadora del Departamento de Derecho UAM-A.

** Profesor-Investigador del Departamento de Sistemas UAM-A.

*** Profesor-Investigador del Departamento de Derecho UAM-A.

Sumario: Introducción / I. La justificación en la decisión judicial / II. La inteligencia artificial generativa en la justificación de la decisión judicial / III. Conclusiones / Fuentes de consulta

Introducción

En la argumentación jurídica resulta indispensable el razonamiento probatorio para acreditar o desacreditar un hecho o actos jurídicos, a través de las evidencias para generar una justificación y convencer al juzgador para obtener una ratio decisión.

Para ello, analizamos la justificación tanto interna como externa desde la perspectiva del jurista italiano Taruffo, con la finalidad de descubrir que elementos pueden tener problemas jurídicos en la aplicación de la fundamentación de las decisiones judiciales con las pruebas electrónicas o científicas, en las cuáles se utilizó indebidamente la Inteligencia Artificial Generativa (IA generativa) con las que se alteró o manipuló la realidad, a través de una imagen, un audio o un video, un texto entre los principales contenidos, representan nuevos desafíos para al juzgador y al perito forense; el primero porque tendrá que justificar su decisión valorando el elemento tecnológico en el dictamen pericial junto a los hechos probados, con todas las demás pruebas, y el segundo porque sólo cuenta con determinadas herramientas para detectar alguna creación, manipulación, o bien, alteración de la realidad en nuestro Derecho, esto debido a que la IA generativa, que está en constante innovación, aunado a lo costoso que resulta crear herramientas para detectar el *deepfake*.¹

Por tales razones, será necesario establecer nuevas bases para futuras teorías de la prueba digital, lo anterior, debido a que encontramos algunas categorías, elementos, características y principios especiales en la realidad virtual, los cuales no se encuentran incluidos en el contexto tradicional y deberán ser incorporadas las decisiones judiciales.

¹ *Deepfake*, que consisten en crear imágenes convincentes, aunque completamente falsas. Por medio del aprendizaje automático de la inteligencia artificial, se manipulan y generan nuevas imágenes o vídeos a partir de otros y se reemplaza a la persona que aparece en ellos. El objetivo es generar diferentes imágenes estáticas para crear una secuencia de video, de modo que, como objetivo final, se obtenga un vídeo falso que parezca 100% real. Dirección de cómputo y comunicaciones, “*Deepfake*. Seguridad de la información”.

Ahora, tenemos pruebas como la documental electrónica que pueden utilizarse simultáneamente, no solo un acto jurídico sino en diversos al mismo tiempo, en las cuales no pueden verse claramente la diferencia entre una prueba original y una copia, excepto con algún peritaje forense en materia de Inteligencia Artificial (IA); además, hoy en día la IA generativa es capaz de reproducir imágenes, diálogos, voces, realizar trabajos de cualquier índole entre muchas de sus capacidades, e incluso tomar algunas decisiones judiciales fáciles.

Resulta necesario determinar, ¿Qué criterios mínimos de certidumbre deberá considerar un juzgador para su decisión judicial fundada y motivada en lo válido, verdadero y justo en el elemento probatorio?, ¿Cómo trasciende el uso de la IA generativa como prueba documental electrónica en la decisión judicial?

Para ello, iniciamos con el autor Taruffo, quien expone los elementos de la justificación, tanto interna como externa, destacando los elementos empíricos que deberán comprobarse para poder ser valorados correctamente, a través de parámetros de valoración de la prueba. Explicaremos cómo hoy en día en las decisiones judiciales el justificar externamente los casos difíciles es más complejo que anteriormente, en virtud de la evolución de las tecnologías emergentes que actualmente nos proporcionan una variedad de soportes electrónicos, digitales, ópticos y de cualquier otra tecnología. También contamos con una variedad de contenidos, de mensajes, e informaciones que ahora son más difíciles de distinguir entre los verdaderos y los falsos por los avances de la IA generativa, en los cuales el juzgador tendrá que incluir el elemento tecnológico para la valoración de la prueba digital, y considerar los dictámenes de expertos en sistemas inteligentes.

Además, analizamos qué es la IA generativa y sus niveles de Reglamento sobre la Inteligencia Artificial del Parlamento Europeo. También examinamos las nuevas problemáticas que se originan con el uso indebido de la IA generativa la cual proporciona datos, información, algoritmos y medios probatorios los cuales son idénticos a los originales, en algunas aplicaciones de IA generativa encontraremos software detectores de la utilización de la IA, pero en algunas otras no habrá programa de cómputo capaz de detectar de dos versiones de textos, imágenes, audios y videos, para saber ¿cuál es el original, y cuál contiene contenido erróneo, falso, manipulado o alterado?, lo cual puede originar diversos problemas en la valoración de las prueba documental

electrónica y la prueba aportada por los avances científicos. Esto debido a que el mundo real está siendo alterado y modificado por la inteligencia artificial generativa que nos presenta una simulación de realidad, dando origen a lo sintético o artificial.

Es necesario la regulación de la IA generativa, estableciendo los elementos normativos tecnológicos, además de “(...) la información pública sobre el uso de los algoritmos; cómo garantizar que la toma de decisiones sea justa e imparcial; y la importancia de la transparencia, la explicabilidad (*sic*) y la confianza”.² Elementos que ayudarán en un futuro a los juzgadores en la elaboración de su justificación para una decisión judicial justa.

I. La justificación en la decisión judicial

El jurista Taruffo, por justificación judicial, indica que es un proceso argumentativo a través del cual el juez explica de manera clara y convincente las razones que lo llevaron adoptar una determinada decisión, y entiende por justificación interna la enunciación de las premisas con base en las cuales se formula una decisión judicial; aquélla es siempre indispensable dado que cualquier aserción es justificable sólo sobre la base de las premisas de las cuales se deriva:

El mundo es mucho más complejo de lo que piensan los lógicos. Ese discurso hecho a nivel de decisiones judiciales no está dicho que sea tan negativo, como cuando se hace a nivel de normas, porque, a nivel de las decisiones judiciales, es donde podemos encontrar la variedad. Encontramos la variedad de la realidad, esto es, el hecho de que ningún caso es exactamente idéntico a otro caso, entonces se podría razonar con una imagen del ordenamiento de varias capas y que, por ejemplo, sea perfectamente coherente y carente de contradicciones sobre el plano de las normas generales, pero en cambio, muy variado y variable en el tiempo y diversificado en la solución de los conflictos individuales y concretos.³

² Estudios de la OCDE sobre Gobernanza Pública, Organización para la Cooperación y el Desarrollo Económico, *Uso estratégico y responsable de la inteligencia artificial en el sector público de América Latina y el Caribe*.

³ Michelle Taruffo, *Cinco lecciones mexicanas: memoria del taller de derecho procesal*, p. 57.

Coincidimos con Taruffo en que el mundo es más complejo, y a nivel de las decisiones judiciales podemos encontrar una variedad, los hechos nunca serán idénticos a otro caso, consideramos que en la actualidad la variedad de la realidad real nos expone nuevos hechos complicados en las decisiones judiciales, las cuales van variando en el tiempo.

Observamos que en la decisiones judiciales surge otro elemento, el tecnológico; debido a la creación de las tecnologías emergentes, las cuales son diversas y dan origen a otras realidades como las llamadas virtual artificiales o sintéticas, y las referidas, crean ficciones y falsedades en nuestro mundo real, como por ejemplo: una simulación de un hecho en el cual se graba la voz de una persona, y se emita un mensaje con esa voz para realizar un hecho ilícito; u otra con la imagen de una persona se alteró o modificó en una foto o un video, para realizar una conducta ilícita; lo cual en un primer momento se consideran hechos verdaderos a simple vista, o bien escuchando al individuo se considera que la voz es la misma. Para determinar si es una realidad real o artificial o sintética será necesario verificar si pertenece al mundo real a través de un peritaje forense.

La prueba pericial demostrará que “la averiguación de la verdad de los hechos es condición necesaria para la justicia de la decisión”.⁴ El razonamiento probatorio debe ser necesario para integrar la decisión judicial, por lo que el peritaje forense se convierte en el elemento tecnológico que se requiere para la comprobación de este hecho que se originó altero o modifíco con la IA generativa, la cual deberá contar criterios mínimos de certidumbre.

Por otro lado, respecto a la justificación externa, Taruffo establece que radica en la convalidación racional de las premisas, es decir, la indicación de las razones por las cuales ellas son asumidas como buenas.

“Apunta que, si bien el uso de métodos científicos para la averiguación de hechos no es nuevo, su creciente adopción sí lo es, fenómeno al que define como ‘el regreso de la verdad’. Celebra el hecho de que en esta tendencia se privilegie una visión multidisciplinaria”,⁵ actuaciones judiciales que en esa medida juegan todos los expertos, es decir, los peritos que proporcionan los elementos científicos, tecnológicos, artísticos o de oficios a través de los dictámenes periciales, el juzgador puede buscar la verdad para emitir su decisión

⁴ Michele Taruffo, *Verdad, prueba y motivación en la decisión sobre los hechos*, p. 13.

⁵ *Ibid.*, p. 10.

judicial, con una visión multidisciplinaria que aporta más elementos de convicción.

Por lo referido, resultarán de gran trascendencia las controversias en las cuales se ofrezcan pruebas documentales electrónicas y peritajes en sistemas inteligentes, los cuales pertenecen a áreas multidisciplinarias de carácter científico-tecnológico, dado que se van a valorar los medios tecnológicos, el tipo de IA generativa que se utilizó.

Por otro lado, “Las otras teorías de la decisión, como la silogística, siempre sirvieron para esconder la responsabilidad del juez y eso quiere decir hacer de él un aplicador pasivo de las directrices del poder, hacer de él alguien que no tiene responsabilidad sino un mero ejecutor, lo que quiere decir es que el poder viene de otra parte”.⁶

Coincidimos en parte con el autor, toda vez que tratándose de casos fáciles el silogismo jurídico sirven para —esconder la responsabilidad— del juzgador, y actúa como un simple ejecutor. Esto no puede ser aplicado en los casos difíciles o complejos que hoy se multiplican ya que, estos últimos requieren de una argumentación “(...) de la racionalidad del razonamiento jurídico donde se entrelazan la razón teórica que diferencia lo verdadero y lo falso, y la razón práctica que trata con lo que debo hacer como individuo (moralmente) o lo que debemos hacer como comunidad (políticamente), pero también con lo que debemos hacer en el plano altamente institucional del derecho”.⁷ “Es preciso incidir en que este acceso al lenguaje no supone que la máquina lo entienda: la Inteligencia Artificial (IA) manipula los símbolos y obtiene correlaciones sintácticas, pero no aprehende ni crea sentido a través de ellos”.⁸ En la valoración de los peritajes en materia de sistemas inteligentes, los juzgadores se enfrentarán con los sistemas inteligentes (grandes modelos de lenguaje), y las decisiones judiciales requieren del elemento científico tecnológico, que deberá ser valorado de acuerdo a los criterios mínimos de certidumbre, de fiabilidad para tener pruebas plenas que le den certeza jurídica al juzgador.

La estructura dialéctica no solo se da entre tesis y antítesis, sino también entre norma y hecho sino por la propia relación procesal: “Entonces, la ‘rela-

⁶ Michele Taruffo, *Cinco lecciones mexicana: memoria del taller de Derecho Procesal*, p. 10.

⁷ Alfonso García Figueroa, *Teoría de la argumentación. Funciones, fines y expectativas*, p. 87.

⁸ Niklas Luhmann, *La sociedad de la sociedad*, p. 413.

ción' es doble (la dialéctica es una cuestión muy complicada). Hay una doble relación entre el hecho y la norma y también entre las partes, cada una de las cuales va a sostener una distinta relación entre ese hecho y esa norma".⁹ El análisis de la tesis y la antítesis, resulta ser complejo, debido al elemento tecnológico. "El juez está en la encrucijada de todas estas relaciones dialécticas y va a ser el que tiene que dar el último paso dialéctico, y decir si reconstruye el hecho de una manera tal e interpreta la norma de esta otra manera. Así es como se verifica la correspondencia".¹⁰ Y esta tendrá hechos modificados o alterados con sistemas de inteligencia artificial, como puede ser un video, una imagen o la voz; y dependerá del tipo de sistema inteligente que se haya utilizado para verificar su certidumbre, porque no todos los sistemas inteligentes tienen aplicaciones que detecten la utilización de la IA generativa, o bien de las técnicas y métodos que puedan detectar el resultado de que ha sido generado o modificado por dicho sistema.

La actividad argumentativa requiere de la capacidad intelectual de los seres humanos emitiendo argumentos racionales que funden y motiven las decisiones judiciales, los cuales estarán sometidos a criterios de racionalidad, por ello, se cuestiona en la actualidad que los sistemas inteligentes realicen una actividad argumentativa, porque se dice que no podrán aportar interpretaciones en los casos complejos. Sí bien la existencia de cajas negras que logrará emular el razonamiento jurídico, no es lo suficientemente racional, porque todo inicia con la obtención de los datos que se incorporan a un sistema, para constituir posteriormente los algoritmos que emularan los razonamientos jurídicos; pero esto representa otro tipo de problema jurídico, como un elemento subjetivo y tecnológico, como lo son los sesgos cognitivos del desarrollador del sistema inteligente, de la identidad del perito, de su independencia, la calidad de las bases de datos con las que se alimenta la IA generativa entre los relevantes.

Por lo que en la justificación se deberá buscar la certeza de todos los elementos, entre ellos el científico-tecnológico, para darle a la prueba pleno valor probatorio o bien, para desacreditarla.

⁹ Michelle Taruffo, *op. cit.*, p. 17.

¹⁰ *Ibid.*

II. *La inteligencia artificial generativa en la justificación de la decisión judicial*

“La IA generativa usa modelos básicos, es decir, grandes modelos de IA, que pueden realizar varias tareas a la vez y realizar tareas listas para usar, como resúmenes, preguntas y respuestas, clasificación, etcétera. Además, al llevar a cabo una preparación mínima, los modelos de base se pueden adaptar a casos prácticos concretos con muy pocos datos de ejemplo”.¹¹

Con el desarrollo de los sistemas de IA generativa se ha incrementado su implementación y uso, en “(...) México el 44% de las compañías empieza a implementar el uso de inteligencia artificial generativa, en sus procesos (...)”.¹² De acuerdo al informe del Foro Económico Mundial en 2023, los videos *deepfake* en internet aumentan a un ritmo anual del 900%.¹³ Y por ello, deberán ser valorados ya sea como prueba electrónica, digital o científica o cualquier otra prueba que proporcione información generada o almacenada en algún soporte informático, digital, óptico o de cualquier otra tecnología:

(...) resulta innegable que la Inteligencia artificial se está impulsando gracias a los avances en algoritmos, minería e datos y *Big Data*, es decir, en la mejora de los procesos de toma de decisiones (los algoritmos) en la mejora de las estructuras de datos y su acceso (minería de datos) y en el acceso a grandes repositorios de información (*Big Data*). Este salto cualitativo es evidente y afectará y de hecho ya está afectando, a prácticamente todos los aspectos personales y profesionales de nuestra vida.¹⁴

La IA generativa puede crear o modificar contenido, como texto, imágenes, música, audio y videos, en estos medios probatorios enfocaremos la justificación en la decisión judicial esto debido a que el perito forense en materia de IA necesita analizar los referidos medios de prueba creados, modificados o alterados con las referidas máquinas inteligentes, para detectar entre dos ver-

¹¹ Google Cloud, “Genera texto, imágenes, código y más con la IA de Google Cloud”.

¹² Mónica Valladolid, “Estas son las barreras de empresas en México ante la inteligencia artificial generativa”.

¹³ Mario Micucci, “Herramientas para detectar *deepfakes* y combatir la desinformación”.

¹⁴ Ejecutivos, “Desafíos de la Inteligencia Artificial en el peritaje informático y análisis de evidencia digital”.

siones ¿cuál es el original, y cuál contiene contenido erróneo o falso, creado, manipulado, alterado por IA generativa?

En algunas ocasiones los peritos cuentan con las aplicaciones o los programas detectores de la utilización de la IA generativa, y se obtendrá certeza jurídica; en otras ocasiones no, consideramos que los altos costos son un gran obstáculo para el desarrollo de las herramientas tecnológicas para detectar *deepfake*, así como por las constantes innovaciones de la IA generativa, dado que el 56% de las empresas afirman que la “imprecisión” es el mayor riesgo que plantea la adopción de la IA generativa “sin embargo, sólo el 32% de las empresas dispone de sistemas para mitigar esas imprecisiones”.¹⁵

Por ello, los dictámenes periciales van a depender mucho del tipo de IA generativa que se utilice para crear el resultado, porque si es una, que no cuenta con los programas detectores será imposible determinar quién realmente realizó o no, el hecho o acto jurídico, lo que podrá originar injusticias dado la nula certeza jurídica que puede originarse con este tipo de pruebas, por los que el elemento científico tecnológico varían de acuerdo al tipo de inteligencia que se utilice, y la confiabilidad deberá tasarse en función de sus particularidades propias de cada sistema inteligente:

El uso de la tecnologías de la Inteligencia Artificial Generativa, ha originado las *deepfakes* (medios sintéticos generados mediante técnicas de aprendizaje profundo que utilizan redes neuronales profundas), es decir, pueden ser imágenes, videos y grabaciones de audio manipulados para retratar a alguien diciendo o haciendo algo que en realidad nunca dijo, ni hizo.¹⁶

Lo referido puede generar incertidumbre e inseguridad jurídica para todos los usuarios de estos sistemas inteligentes, para lo cual se requiere formular estándares técnicos y tecnológicos que determinen los requisitos mínimos para darle valor probatorio al elemento científico tecnológico, se establezcan criterios de certidumbre y seguridad informática con la finalidad de prevenir riesgos específicos de suplantación o engaño, con independencia de si cumplen las condiciones para ser considerados como de alto riesgo o no.

¹⁵ *Ibid.*

¹⁶ Mario Micucci, “Herramientas para detectar *deepfakes* y combatir la desinformación”.

Por consiguiente, el uso de estos sistemas debe estar sujeto, en determinadas circunstancias, a obligaciones de transparencia específicas, sin perjuicio de los requisitos y las obligaciones aplicables a los sistemas de IA de alto riesgo y a excepciones específicas a fin de tener en cuenta las necesidades especiales de la garantía del cumplimiento del Derecho.¹⁷

En México actualmente no se tiene una regulación de la IA, ni mucho menos se tiene una clasificación de los riesgos de ésta, ni se prohíbe su uso, lo cual originará otros tipos de problemas jurídicos para el juzgador en el análisis de la valoración de la prueba electrónica o científica, se deberá examinar el elemento científico-tecnológico, que aparecerá en el razonamiento probatorio, porque el juzgador se encontrará en dicho elementos tanto subjetivos como nuevos problemas, como: a) los sesgos cognitivos del programador y del perito, b) la calidad de los datos, c) la falta de consentimiento en la utilización de base de datos para sistemas inteligentes, d) la falta de ética de los desarrolladores, e) la falta de estándares de seguridad informática que ponen en duda y controversia las pruebas obtenidas con el uso de sistemas inteligentes, f) la transparencia de los desarrolladores, g) la calidad e imparcialidad del perito, i) la falta de técnicas y métodos para detectar algunas *deepfake*, j) la confiabilidad o fiabilidad de la prueba, k) los estándares de calidad de los sistemas inteligentes, entre los principales.

En México será necesario establecer reglas de transparencia para obligar a los programadores de IA generativa que hagan de conocimiento a los usuarios de la IAG, que cuentan con programas detectores de IA generativa, o no, esto con la finalidad de que el usuario acepte o no, el riesgo que corre con la implementación de estas tecnologías, además de contar con los mínimos de racionalidad para poder determinar si puede ser o no justa la decisión judicial.

II.1. ¿Qué es la inteligencia artificial generativa?

El Reglamento sobre Inteligencia Artificial promulgado el 13 de marzo de 2024 por el Parlamento Europeo,¹⁸ ha definido al sistema de IA (art. 3,1) como un sistema basado en una máquina que está diseñado para funcionar con distintos niveles de autonomía y que puede mostrar capacidad de adaptación tras el despliegue (puesta en funcionamiento), y que, para objetivos explícitos

¹⁷ Parlamento Europeo, *Reglamento de Inteligencia Artificial*.

¹⁸ Parlamento Europeo, *Artificial Intelligence Act*.

o implícitos, infiere de la información de entrada que recibe la manera de generar resultados de salida, como predicciones, contenidos, recomendaciones o decisiones, que pueden influir en entornos físicos o virtuales.¹⁹ Lo anterior ha sido duramente criticado por la Cámara de Comercio Americana en la UE²⁰ por ser una definición demasiado extensiva que abarca a los sistemas inteligentes de bajo riesgo, lo cual puede originar en Europa una excesiva regulación.

Es importante indicar que de acuerdo a un análisis de la compañía de soluciones de datos SAS, el 92% de los ejecutivos responsables de la tecnología en su organización reconoce no entender qué es la Inteligencia Generativa, aseguró la empresa tras haber hecho una consulta.²¹ El desconocimiento y la desinformación de lo que son las inteligencias artificiales, son factores que propician las vulneraciones a derechos fundamentales de los sujetos que interactúan en los espacios virtuales.

II.2. Clasificación de inteligencias: riesgosas

Como toda tecnología, la IA representa una herramienta que debe estar centrada en el usuario, el ser humano. Sin embargo, como toda herramienta puede dar lugar a fines positivos o negativos, desde el punto de vista del bienestar humano. En este sentido, aunado a los estándares de seguridad informática, es relevante el avance del Reglamento de Inteligencia Artificial del Parlamento Europeo, que clasifica los sistemas de IA en: sistemas de IA de alto riesgo, sistemas de IA de riesgo limitado, y prácticas de IA prohibidas.

- a) Sistemas de IA de alto riesgo: Pueden tener un alto grado de influencia en intereses públicos de salud, seguridad y de los derechos fundamentales.
- b) Sistemas de IA de riesgo limitado: No influyen sustancialmente en el resultado de la toma de decisiones, y está destinada a mejorar el resultado de una actividad previa realizada por un humano.
- c) Prácticas IA prohibidas: Que van en contra de la dignidad humana, la libertad la igualdad, la democracia y el Estado de derecho.

¹⁹ *Ibid.*

²⁰ *Ibid.*

²¹ Valladolid, *op. cit.*

Los sistemas de IA se diferencian de otros sistemas de cómputo tradicionales en su capacidad de inferencia mediante tareas de predicción, recomendación o toma de decisiones con base a datos de entrada, a partir de los cuales puede deducir modelos o algoritmos que aluden al aprendizaje automático toda vez que logran emular sistemas de razonamiento o aprendizaje con cierto grado de independencia respecto a la participación humana.

Los modelos de IA de uso general, aprenden a partir de grandes cantidades de datos, como es el enfoque de IA generativa, el cual tiene un grado mínimo de intervención humana. Por ejemplo, a partir del concepto de coevolución, se hace competir a dos sistemas antagónicos, uno que propone (generador) y otro que evalúa (discriminador), de forma tal que en un principio el primer sistema genera datos de salida con gran aleatoriedad a partir de datos de entrenamiento, con el objetivo de que las siguientes salidas capten los patrones de los datos de entrenamiento sin que coincidan con ellos, de lo cual se encarga de evaluar el segundo sistema.

Al continuar con el proceso, ambos sistemas se perfeccionan para producir datos de salida con características comunes a los datos de entrenamiento, de forma tal que al evaluador se le dificulte identificar si el dato proviene del conjunto de entrenamiento o ha sido generado por su contraparte, el sistema generador, y entonces se ha construido un sistema artificial auto generativo.

Otros modelos de IA de uso general se entrenan a partir de grandes cantidades de texto, en donde se logran identificar reglas gramaticales, correlaciones semánticas, no solo a nivel de palabras sino de frases completas, con alto grado de coherencia.

En este orden de ideas, es preciso indicar con respecto a las *deepfake*, en el Reglamento sobre Inteligencia artificial en el art. 50 numeral 2, se regula lo siguiente:

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 particulari-

dades 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.²²

De lo que se desprende que los proveedores de sistemas de IA de uso general que produzcan 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 este contenido fue generado o manipulado de manera artificial, como los requisitos precautorio de seguridad, es ésta obligación ayudaría a los peritos forenses en materia de IA a determinar la integridad, fiabilidad y transparencia de las pruebas ofrecidas en los procesos judiciales, en el caso de que se hubiese usado indebidamente la IA generativa.

II.3. El funcionamiento de la Inteligencia artificial generativa, algunos usos indebidos y la necesidad de su regulación

Por otro parte, el funcionamiento de la IA generativa usa un modelo de aprendizaje automático para aprender los patrones y las relaciones de un conjunto de datos de contenido creado por personas. A continuación, utiliza los patrones aprendidos para generar contenido. La forma más habitual de preparar un modelo de IA generativo es usar el aprendizaje supervisado, que es un conjunto de contenido creado por humanos y las etiquetas correspondientes. A continuación, aprende a generar contenido similar al creado por personas y etiquetado con las mismas etiquetas.²³

Es importante indicar que la IA generativa tiene una variedad de aplicaciones entre las que consideramos:

- a) Los textos/chatbots, son herramientas que trabajan con textos e información, principalmente a través de asistentes virtuales conocidos como chatbots.

²² Parlamento Europeo, *Reglamento de Inteligencia Artificial*.

²³ *Ibid.*

- b) Video-Herramientas multimedia que generan o editan videos de forma automatizada.
- c) Imágenes. Son herramientas multimedia que generan o editan dibujos o fotografías de forma automatizada.
- d) Aplicaciones de inteligencia artificial que generan música o archivos de voz (*tex-to-speech* y *speech to tex*), el uso indebido de algunas de estas aplicaciones generarán actos y hechos ilícitos, delitos y conductas ilícitas.

Las prácticas perjudiciales que originan la diversidad de aplicaciones de IA generativa, tendrán como resultado una diversidad de controversias y el problema de saber ¿si podrá verificarse la originalidad de un contenido de audio, imagen, vídeo o texto de con las tecnologías de IA?

Esto implica tener en consideración que el juzgador al justificar su decisión judicial, deberá considerar tanto la norma jurídica, como la norma tecnológica, los hechos y las pruebas, entre las cuales el peritaje tecnológico o forense en IA va a tener obstáculos para poder buscar herramientas tecnológicas para detectar el engaño o falsedad en algunas aplicaciones como lo es el reconocimiento de la voz, pues en estos momentos no existe software que permita determinar las frecuencias de la voz sintética de la voz humana fehacientemente, por lo que el juzgador no podrá tener la certeza jurídica de determinar de quién es la voz que le presentan en un soporte electrónico, porque la voz clonada o artificial, se obtuvo de los datos de entrenamiento.

La variedad de aplicaciones de IA generativas, así como la falta de programas que pueda detectar las modificaciones o engaños causados por el uso indebido de aplicaciones de IA generativa, podría generar que no se realice una debida justificación en la decisión judicial, porque el juzgador no tendrá los elementos normativos tecnológicos para poder determinar si hubo o no alguna modificación, alteración, suplantación, o reemplazo de algún video, texto, imagen o audio; en la actualidad las *deepfakes*, además de desinformar a los cibernautas, también generan delitos, conductas ilícitas, actos y hechos ilícitos que vulneran los derechos fundamentales.

Ahora bien, en algunos sistemas de IA, ya se han prohibido prácticas perjudiciales e incorrectas en Europa, en virtud de que vulneran derechos fundamentales, y representan riesgos graves a los usuarios y demás sujetos en el espacio virtual, el “(..) ochenta por ciento de los EE. UU. a los adul-

tos les preocupa que la IA se utilice para ciberataques. Y al 78% le preocupa que se utilice para robar identidades. (Mitre, 2023).²⁴ En la Ciudad de México de acuerdo a las cifras del Consejo Ciudadano para la Seguridad y Justicia de la Ciudad de México revelan que, en el último año, “la usurpación de identidad con Inteligencia Artificial se incrementó en un 218 por ciento”,²⁵ lo cual nos refleja el aumento de vulneraciones de derechos fundamentales con el uso de IA.

En Europa antes de la creación del Reglamento sobre Inteligencia Artificial, no se regulaban las obligaciones de transparencia, lo cual ocasionaba una vulneración de los derechos fundamentales de los usuarios, dado que el 56% de las empresas afirman que la “imprecisión” es el mayor riesgo que plantea la adopción de la IA generativa “sin embargo, sólo el 32% de las empresas dispone de sistemas para mitigar esas imprecisiones”.²⁶

En el Reglamento sobre Inteligencia Artificial del Parlamento Europeo, se regulan las obligaciones para los implementadores de sistemas de IA, por lo que consideramos que el referido reglamento contiene la mayoría de soluciones como por ejemplo las siguientes:

- a) Los implementadores de sistemas de IA que generen o manipulen contenidos de imagen, audio o vídeo (es decir, *deepfakes*) deben revelar que el contenido ha sido generado o manipulado artificialmente, excepto en casos muy limitados (por ejemplo, cuando se utiliza para prevenir delitos). Es decir, se establecen obligaciones de transparencia para los implementadores de sistema de IA, lo cual dará certeza a los usuarios.
- b) Los proveedores de sistemas de IA que crean grandes cantidades de contenido sintético o simulado deben aplicar técnicas y métodos suficientemente fiables, interoperables, eficaces y robustos (como las marcas de agua) para permitir el marcado y la detección de que el resultado ha sido generado o manipulado por un sistema de IA y no por un humano.

Consideramos que esta es una de las obligaciones para los proveedores de IA que generan grandes cantidades de contenido; que puede aportar más elementos científicos técnicos y tecnológicos para la justificación

²⁴ Padrig Jones, “78 estadísticas y Tendencias de la Inteligencia Artificial para 2024”.

²⁵ INAI, Comunicado 373/23.

²⁶ *Ibid.*

en la decisión judicial, porque se establecen obligaciones para los proveedores, implementar técnicas y métodos fiables que permiten el marcado y la detección de que el resultado ha sido generado o manipulado por IA generativa, elemento científico tecnológico que en México actualmente no es una obligación, y que será un obstáculo, para los peritos forense y para los juzgadores.

- c) Los empleadores que implementen sistemas de IA en el lugar de trabajo deben informar a los trabajadores y sus representantes.

Por lo que sí el Estado mexicano adopta esta regulación estaremos solucionando la mayoría de problemas respecto al razonamiento probatorio, y la justificación de la decisión judicial, pues el juzgador contará con más elementos para analizar si el contenido ha sido resultado de una generación o manipulación de un sistema de inteligencia generativa, o bien si la imagen, el audio, el video o texto han sido generados por dicho sistema, lo cual permitirá al juzgador tener más elementos probatorios para poder justificar y determinar una decisión judicial justa.

III. Conclusiones

Primera. En virtud de la diversidad de las inteligencias artificiales, es necesario que se regule la IA general y generativa, considerando los riesgos que ocasionan, como ya se ha regulado en Europa, a través de una clasificación de acuerdo a los riesgos, lo cual puede ser un modelo a seguir para una posterior regulación en México.

Segunda. El peritaje forense en materia de IA será el elemento normativo-tecnológico que deberá considerar el juzgador de manera integral, para realizar su decisión judicial, pero será necesario establecer en nuestra regulación mexicana, las obligaciones que tendrán tanto los desarrolladores y los proveedores de IA generativa, que contribuirán a toma de decisiones judiciales más justas.

Tercera: En México sólo existen algunas aplicaciones detectoras para la distinción entre lo real y la realidad artificial o sintética como resultado de generado o manipulado por un sistema de IA y no por un humano.

Fuentes de Consulta

Bibliográficas

- García Figueroa, Alfonso. *Teoría de la argumentación. Funciones, fines y expectativas. Argumentación Jurídica*. Marina Gascón Abellán y Alfonso García Figueroa (eds.), España, Tirant lo Blanch, 2014.
- Luhmann, Niklas. *La sociedad de la sociedad*. J. Torres Nafarrate, (trad.), México, Herder, 2007.
- Taruffo, Michelle. *Verdad, prueba y motivación en la decisión sobre los hechos*. México, Tribunal Electoral del Poder Judicial de la Federación, 2013.
- _____. *Cinco lecciones mexicanas: Michele Taruffo, memoria del taller de Derecho Procesal*. México, Tribunal Electoral del Poder Judicial de la Federación, 2003.

Electrónicas

- Estudios de la OCDE sobre Gobernanza Pública. Organización para la Cooperación y el Desarrollo Económico. Uso estratégico y responsable de la inteligencia artificial en el sector público de América Latina y el Caribe, OCDE. file:///C:/Users/alega/Downloads/5b189cb4-es.pdf
- Ejecutivos. “Desafíos de la Inteligencia Artificial en el peritaje informático y análisis de evidencia digital”. España, 2023. <https://ejecutivos.es/reportaje/desafios-de-la-inteligencia-artificial-en-el-peritaje-informatico-y-analisis-de-evidencia-digital/> (Consultada el 8 de septiembre de 2024).
- Google cloud. “Genera texto, imágenes, código y más con la IA de Google Cloud”. <https://cloud.google.com/use-cases/generative-ai?hl=es> (Consultada el 13 de septiembre de 2024).
- Jones, Padrig. “78 estadísticas y tendencias de la Inteligencia Artificial para 2024”. <https://es.semrush.com/blog/etica-de-la-inteligencia-artificial/> (Consultada el 2 de septiembre).
- Micucci, Mario. “Herramientas para detectar deepfakes y combatir la desinformación”. <https://www.welivesecurity.com/es/seguridad-digital/herramientas-para-detectar-deepfakes-combatir-desinformacion/> (Consultada el 27 de septiembre de 2024).
- Parlamento Europeo. *Artificial Intelligence Act*. [https://www.europarl.europa.eu/think-tank/es/document/EPRS_BRI\(2021\)698792](https://www.europarl.europa.eu/think-tank/es/document/EPRS_BRI(2021)698792) (Consultada el 2 de septiembre de 2024).
- _____. Reglamento de Inteligencia Artificial. <https://data.consilium.europa.eu/doc/document/PE-24-2024-INIT/es/pdf> (Consultada el 2 de septiembre de 2024).
- Valladolid, Mónica. “Estas son las barreras de empresas de México ante la inteligencia artificial generativa”. México, Forbes, 2024. <https://www.forbes.com.mx/estas-son-las-barreras-de-empresas-en-mexico-ante-la-inteligencia-artificial-generativa/> (Consultada el 30 de septiembre de 2024).

Hemerográficas

Dirección de Computo y Comunicaciones. “*Deepfake*. Seguridad a la Información”. Secretaría de Educación Pública, Instituto Politécnico Nacional. https://www.seguridad.ipn.mx/comunicados/Infografia_Deepfake.pdf

Instituto Nacional de Transparencia, Acceso a la información y Protección de datos personales. “Comunicado 373/23”. <https://home.inai.org.mx/wp-content/documentos/SalaDePrensa/Comunicados/Comunicado%20INAI-373-23.pdf> (Consultado el 21 de septiembre de 2024).