

Notas sobre el uso de la IA generativa por profesionales del sistema de justicia

Joaquín Delgado Martín

Magistrado Sala Penal Audiencia Nacional

Miembro de la Red de Especialistas en Derecho UE (REDUE)

Doctor en Derecho

Diario LA LEY, Nº 10568, Sección Tribuna, 16 de Septiembre de 2024, LA LEY

ÍNDICE

[Notas sobre el uso de la IA generativa por profesionales del sistema de justicia](#)

[I. ¿Qué es la IA generativa?](#)

[II. Uso de la IA generativa por profesionales del sistema de justicia](#)

[III. Riesgos](#)

[IV. Uso de ChatGPT 3.5 por un juez: sentencia de la Corte Constitucional de Colombia](#)

Comentarios

Resumen

Tomando como punto de partida los datos de la utilización creciente por los profesionales de justicia de la IA generativa, el autor analiza tanto la utilidad de su uso, como los riesgos inherentes al mismo.

I. ¿Qué es la IA generativa?

La IA generativa permite a un usuario crear contenidos originales (como texto, imágenes, vídeo, audio o código de *software*) en respuesta a sus peticiones. Se fundamenta en complejos modelos de machine learning llamados de deep learning, es decir, algoritmos que simulan los procesos de aprendizaje y toma de decisiones del cerebro humano. Este tipo de modelos funcionan identificando y codificando los patrones y las relaciones en enormes cantidades de datos, de tal forma que posteriormente utiliza esa información para entender las peticiones o preguntas en lenguaje natural de los usuarios, respondiendo con nuevos contenidos relevantes. Su utilización ha sufrido un gran impulso desde la introducción en el mercado de ChatGPT en 2022 (OpenAI), habiendo asistido a los lanzamientos de otros productos como Google Bard (ahora Gemini), Microsoft Copilot, IBM Watsonx.ai ó Llama-2 de Meta.

Chat GPT analiza un gran número de palabras para poder predecir la palabra más probable para continuar una frase, y cada una de las palabras siguientes. Hay un gran número de combinaciones posibles de palabras que pueden aparecer a continuación en cualquier frase. Para predecir con fiabilidad una palabra que tenga sentido en una situación determinada, un modelo de IA debe comprender cómo encajan las palabras. Esencialmente, un modelo tiene que aprender todos los elementos del lenguaje: vocabulario, gramática, estructura de las frases e incluso razonamiento básico. Este es el objetivo del proceso de entrenamiento de la IA. El proceso de entrenamiento consiste en dar al modelo una gran cantidad de frases para que las «estudie». El modelo revisa cada frase palabra por palabra e intenta predecir la siguiente palabra de cada frase. Cuando comienza el proceso de entrenamiento, el modelo no sabe cómo predecir palabras y simplemente emite palabras aleatorias. Con suficientes intentos y frases, el modelo aprende con el tiempo mejores palabras que se ajustan a cada frase (1) .



El uso de este tipo de IA se ha generalizado en los diferentes sectores de la sociedad, y múltiples entidades implementado aplicaciones de IA generativa, o bien utilizando interfaces de programación de aplicaciones (API) de IA generativa.

II. Uso de la IA generativa por profesionales del sistema de justicia

Pues bien, la IA generativa también se está utilizando de forma creciente por los profesionales del sistema de justicia. Según la «Iniciativa Global de Jueces de la UNESCO: Encuesta sobre el Uso de Sistemas de IA por Operadores Judiciales» (2024) (2) , el 44% de los operadores judiciales habían utilizado herramientas de IA para actividades relacionadas con el trabajo; y el 41% dijo que habían usado ChatGPT u otros chatbots de IA. Al preguntarles con qué frecuencia utilizan sistemas de IA en su trabajo, el 20% de los encuestados respondió que habían utilizado sistemas de IA solo unas pocas veces, el 13% semanal o mensualmente, y el 11% diariamente. En cuanto a cómo utilizan la información producida por los chatbots de IA, el 55% consideraban el resultado del sistema para redactar sus textos; el 39% usaban el resultado del chatbot pero solo después de realizar un proceso de revisión y edición; y el 6% utilizaban directamente el resultado del chatbot tal cual, sin ninguna revisión o verificación.

Dicho informe constata que los operadores judiciales abogan abrumadoramente por la necesidad de directrices para el uso de herramientas de IA; y que la mayoría de los encuestados creen que debería haber regulaciones obligatorias sobre el uso de chatbots de IA para los diferentes tipos de operadores judiciales. Más específicamente, el 73% de los encuestados consideró que debería haber normas obligatorias para jueces y fiscales, el 72% para abogados y el 73% para funcionarios en la administración legal e investigadores.

En relación con las citadas directrices, cabe destacar la «Artificial Intelligence (AI): Guidance for Judicial Office Holders» (12 diciembre 2023) del Poder Judicial de Inglaterra y Gales (3) . Esta guía ha sido desarrollada para asistir a los titulares de cargos judiciales en relación con el uso de la inteligencia artificial (IA). Establece los riesgos y problemas clave asociados con el uso de la IA y algunas sugerencias para minimizarlos. También incluye ejemplos de usos potenciales. Cualquier uso de la IA por parte del poder judicial o en su nombre debe ser consistente con la obligación general del poder judicial de proteger la integridad de la administración de justicia. Esta guía se aplica a todos los titulares de cargos judiciales, sus secretarios y otro personal de apoyo. Esta Guía destaca los siguientes elementos: comprender la IA y sus aplicaciones; mantener la confidencialidad y la privacidad; garantizar la responsabilidad y la precisión; ser consciente de los sesgos: mantener la seguridad: asumir la responsabilidad; y tener en cuenta que los usuarios del tribunal/juzgado pueden haber utilizado herramientas de IA.

III. Riesgos

El uso de instrumentos de IA generativa en la justicia tiene **riesgos de imprecisión**, es decir, de producir resultados inexactos, incompletos, engañosos o desactualizados. Téngase en cuenta que la salida que generan los chatbots de IA es lo que el modelo predice como la combinación de palabras más probable, basada en los documentos y datos que tiene como información fuente. Téngase en cuenta que, por la forma de funcionamiento de estos sistemas IA, no siempre la siguiente palabra más probable es la más correcta o adecuado desde el punto de vista fáctico y/o jurídico. De esta manera, el sistema IA puede aportar sentencias ficticias; referirse a legislación, artículos o textos legales que no existen; o puede inventarse citas o referencias.

Este tipo de IA puede resultar **útil** para hallar material ya conocido como correcto, pero que no se encuentre fácilmente disponible; para realizar una primera aproximación, pero siempre que después el resultado sea confirmado con fuentes fiables; o bien para resumir grandes volúmenes de texto aunque, como con cualquier resumen, se debe tener cuidado para asegurar que el resumen sea preciso; para escribir presentaciones; y/o para redactar correos electrónicos y diversos tipos de memorándum. Sin embargo, su utilización no resulta aconsejable para realizar investigaciones con la finalidad de encontrar nueva información legal que no se puede verificar.

También es verdad que este tipo de sistemas IA son **«entrenados»** por el usuario, de tal manera que la calidad de las respuestas que reciba dependerá de cómo se interactúe con la herramienta IA, incluyendo la naturaleza de las indicaciones que ingrese. Sin embargo, incluso con las mejores indicaciones, la información proporcionada por el sistema IA puede ser inexacta, incompleta, engañosa o sesgada.

Su utilización también puede determinar problemas sobre la **confidencialidad y protección de datos personales**.

Téngase en cuenta que los chatbots de IA públicos pueden recordar cada pregunta que les hace, así como cualquier otra información que les proporcione; y esa información puede estar disponible posteriormente para responder a consultas de otros usuarios. Por ello, resulta recomendable tratar todas las herramientas de IA como si fueran capaces de hacer público cualquier dato que se ingrese en ellas.

La utilización de instrumentos de IA generativa por personas que realizan actos procesales **sin asistencia letrada** debe realizarse con mayores cautelas, especialmente cuando se trate de la única fuente de asesoramiento o asistencia que reciban la parte procesal. Hay que tener presente que carecen de las habilidades para verificar de manera independiente la información legal proporcionada por los chatbots de IA; y pueden no ser conscientes de los riesgos de imprecisión.

Por último, cabe destacar que el uso de herramientas de IA generativa puede aportar **riesgo de alteración de pruebas** en los procesos judiciales, porque pueden ser utilizadas para producir material falso (textos, imágenes y videos) que sea aportado como prueba a juicio (tecnología deepfake). El sistema de justicia debe ser consciente de esta posibilidad.

IV. Uso de ChatGPT 3.5 por un juez: sentencia de la Corte Constitucional de Colombia

La Sentencia de la Corte Constitucional de Colombia T-323 de 2024 se refiere a un supuesto de aplicación de ChatGPT 3.5 por parte del Juzgado del Circuito en la sentencia de tutela de segunda instancia del 30 de enero de 2023. El juez del Circuito advirtió que haría uso de herramientas de inteligencia artificial generativas para «extender los argumentos de la decisión adoptada», recogiendo las preguntas realizadas al aplicativo y sus respuestas.

La Corte Constitucional considera que «el uso de IA generativa en la administración de justicia, sin salvaguardas, puede afectar el deber de motivar las decisiones judiciales (4) , las garantías del juez natural (5) independiente e imparcial (6) y el derecho a la defensa al producirse desconfianza sobre la fiabilidad de las decisiones y, por tanto, promoverse que fácticamente se dé un reemplazo del juez por parte de la herramienta y la adopción de los fallos dentro del sistema judicial. La protección de estos derechos, se encuentra basada en el cumplimiento del principio de transparencia y explicabilidad por parte de los jueces, ya que se hace imprescindible que las partes interesadas conozcan los fundamentos usados por estos para que en contraposición se puedan tomar las decisiones procesales adecuadas, en particular, frente al ejercicio del derecho de contradicción, de lo contrario, un mal uso de las IA puede conllevar a la vulneración del derecho fundamental al debido proceso» (apartado 273).

(1) Informe utilizado en el caso abordado por la Sentencia de la Corte Constitucional de Colombia T-323 de 2024

(2) <https://unesdoc.unesco.org/ark:/48223/pf0000389786/PDF/389786eng.pdf.multi>

(3) <https://www.judiciary.uk/guidance-and-resources/artificial-intelligence-ai-judicial-guidance>

(4) Considera que «el uso de IA puede arrojar output incorrectos o falsos que de ser usados en la construcción de una decisión judicial pueden llevar a que su motivación sea incorrecta o falsa y, por tanto, inválida (apartado 256).

(5) En relación con la garantía del juez natural, la Corte Constitucional estima que «parte del contenido esencial de la garantía del juez natural es la condición de ser humano que debe asistirle al juzgador» (apartado 213).

(6) Afirma que «...al usarse IA generativas por parte de los jueces, la imparcialidad debería predicarse también de la herramienta pues, dependiendo de sus sesgos, podrían verse afectada algunas de las partes en la causa, con el agravante de que dicha parcialidad puede llegar a ser desconocida o no percibida y pone en entredicho la transparencia de la judicatura» (apartado 263)
