

Título: **El asistente de inteligencia artificial de Microsoft Copilot "declara" en juicio que depende de la supervisión y análisis humanos**
Autor: de la Fuente, María del Rosario
País:  Argentina
Publicación: El Derecho - Diario, Tomo 309
Fecha: 26-11-2024 Cita Digital: ED-V-CMXXII-349

El asistente de inteligencia artificial de Microsoft Copilot "declara" en juicio que depende de la supervisión y análisis humanos

Opinión en El Derecho

por María del Rosario de la Fuente(*)

Sumario: I. Introducción. - II. El uso del asistente de inteligencia artificial de Microsoft Copilot por parte del experto en su segundo informe. - III. El análisis de la prueba por el juez en el asunto de Weber. - IV. Conclusión.

I. Introducción

Excelente noticia para los abogados y los estudiantes de Derecho: la inteligencia artificial "reconoce" sus límites en juicio. Tal reconocimiento ocurrió en el asunto de Weber ante el juez subrogante del condado de Saratoga en el estado de Nueva York (EE. UU.) en octubre pasado(1). El juez Jonathan G. Schopf determinó que los cálculos y el uso de inteligencia artificial (Microsoft Copilot) por parte del experto eran inherentemente poco confiables.

El asunto de Weber comprende la contabilidad y rendición de cuentas del fideicomiso de Michael S. Weber, específicamente, sobre la venta de una propiedad situada en la isla Cat en el archipiélago de las Bahamas en 2022. El demandante y beneficiario del fideicomiso (Owen K. Weber) alegó que la fiduciaria (Susan F. Weber) violó su deber fiduciario al no vender la propiedad con anterioridad a 2022 y al utilizarla para viajes personales. El 10 de octubre pasado, el juez determinó que el demandante no probó que la venta de la propiedad recién en 2022 haya ocasionado daños al fideicomiso. Y decidió que los viajes de la fiduciaria a la propiedad en Bahamas (por motivos de trabajo y por vacaciones personales) no constituyeron violación a su deber fiduciario. También señaló que la evidencia examinada resultó insuficiente para establecer que los viajes de la fiduciaria a Bahamas influenciaron su decisión de no vender la propiedad antes de 2022.

Esta nota tiene por objeto enfatizar que la inteligencia artificial es solo un instrumento y, por ello, es necesario garantizar y proteger un control significativo del ser humano sobre los resultados de su uso. Brevemente, describiremos el uso de Microsoft Copilot por parte del experto en su segundo informe (II) y el análisis de la prueba por el juez (III) en el asunto de Weber.

II. El uso del asistente de inteligencia artificial de Microsoft Copilot por parte del experto en su segundo informe

El Sr. Charles Ranson intervino en el asunto de Weber en calidad de experto en litigios sobre fideicomisos(2). Presentó el cálculo de los daños por el lucro cesante que la venta de la propiedad en Bahamas en 2022 le ocasionó, en su opinión, al fideicomiso. Elaboró dos informes: el informe preliminar de fecha 14 de diciembre de 2022, y el informe complementario de daños de fecha 28 de mayo de 2024, siendo en la producción de este último documento donde empleó al asistente de inteligencia artificial de Microsoft: Copilot.

Con independencia del uso de la inteligencia artificial en el segundo informe, la labor del experto incluyó dos errores que un experto no debe cometer. En primer lugar, aplicó una versión obsoleta de la ley de inversor prudente y, luego, calculó los daños desde la fecha de 31 de diciembre de 2004. Esa fecha es incorrecta, ya que el fideicomiso adquirió la propiedad con fecha 25 de enero de 2008, conforme la decisión judicial sobre la herencia del Sr. Michael S. Weber. Otros errores cometidos por el experto (que no involucran el uso de la inteligencia artificial) fueron: no tener en cuenta en sus cálculos el pago de impuestos por la propiedad en los años 2004-2013(3), 2016-2017, y 2020; no considerar el impacto de la pandemia COVID-19 en los alquileres de las propiedades en el archipiélago de Bahamas en 2020-2021; no incluir en sus cálculos los desembolsos (pagos) al beneficiario del fideicomiso; tampoco considerar las exenciones impositivas y cancelaciones en sus cálculos ni la reinversión del precio obtenido por la venta de la propiedad en 2022.

El experto reconoció no contar con conocimientos sobre el mercado inmobiliario en el archipiélago de las Bahamas ni en otro lugar. Admitió no haber empleado un análisis A. M. R. que constituye el análisis estándar de su campo de especialización(4). El experto testificó que realizar tal análisis implicaba contar con doscientas dieciséis declaraciones de inversión del fideicomiso para obtener los datos sobre los costos de activos, dividendos, intereses, y flujos de dinero, entre otros. Por todos estos errores y omisiones, el juez determinó que el testimonio del experto y sus informes (sin considerar el uso de la inteligencia artificial) resultaban inherentemente poco confiables.

En su testimonio, el Sr. Ranson reveló haber confiado en el asistente de inteligencia artificial de Microsoft (Copilot) para verificar sus cálculos. A pesar de su dependencia en el uso de Copilot, el experto no pudo recordar qué carga de datos (input) o mensaje (prompt) empleó para producir su informe complementario de daños. Tampoco pudo indicar las fuentes en las que Copilot se basó para formular el cálculo ni supo explicar cómo funciona Copilot ni brindar detalles de cómo arriba a un resultado (output).

Copilot es un asistente virtual de inteligencia artificial desarrollado por Microsoft en febrero de 2023. Es un chatbot, es decir, un tipo de inteligencia artificial generativa que es capaz de interactuar directamente con seres humanos y utiliza un modelo de lenguaje de gran tamaño. Permite consultar temas de historia, ciencia, gastronomía, tecnología, cultura general, entre otros, y comprender fotografías y generar imágenes. Está disponible en ocho idiomas (inglés, francés, castellano, alemán, italiano, japonés y chino). Como todo chatbot, está diseñado para aprender a responder -de forma personalizada- a las necesidades físicas y psicológicas de las personas. Responde preguntas, escribe código, crea presentaciones y edita documentos.

El Sr. Ranson destacó, durante su testimonio, que el uso de Copilot u otra herramienta de inteligencia artificial con el fin de redactar informes periciales está generalmente aceptado en el ámbito de los servicios fiduciarios y representa el futuro de las decisiones fiduciarias.

III. El análisis de la prueba por el juez en el asunto de Weber

El magistrado estableció que el testimonio del Sr. Ranson era poco confiable, ya que el experto admitió no haber realizado un análisis completo del problema. También aceptó haber utilizado un período de tiempo incorrecto para el cálculo de los daños, y no haber considerado ciertos rubros básicos en sus cálculos. Todo ello atentó contra su credibilidad como experto.

Sin embargo, el juez reconoció en su fallo que “el uso de la inteligencia artificial es una realidad de rápido crecimiento en muchas industrias”(5). Y agregó que “el mero hecho de que la inteligencia artificial ha jugado un papel, que continúa expandiéndose en nuestra vida cotidiana, no hace que los resultados generados por la inteligencia artificial sean admisibles ante los tribunales”(6).

A pesar de que el testimonio y los informes del experto resultaron poco confiables sin necesidad de examinar el rol que Copilot tuvo en la generación del segundo informe, el juez procedió a comprobar los resultados obtenidos por el experto. Para ello, empleó tres computadoras del juzgado(7). En cada una de ellas, empleando Copilot, preguntó: ¿puedes calcular el valor actual de una inversión por US\$ 250.000 en el fondo de inversión Vanguard Balanced Index desde el 31 de diciembre de 2004 hasta el 31 de enero de 2021? El juez obtuvo tres respuestas distintas, a saber: a) US\$ 949.070,97; b) US\$ 948.209,63; y c) US\$ 951.000. La nimia diferencia entre las respuestas cuestionó la exactitud y la fiabilidad de la prueba generada por Copilot sobre la que podría basarse un pronunciamiento judicial.

El juez indagó más. Preguntó a Copilot si era preciso. Obtuvo la siguiente respuesta: “Debo ser preciso con los datos con los que me han entrenado y la información que puedo encontrar para usted. Dicho esto, mi precisión es tan buena como mis fuentes, por lo que, para temas importantes, siempre es prudente verificar”(8).

El magistrado consultó a Copilot si era confiable. El chatbot contestó: “Por supuesto. Cuando se trata de proporcionar información y participar en conversaciones, hago lo mejor que puedo para ser lo más confiable posible. Sin embargo, estoy programado para aconsejarle que consulte con expertos para asuntos importantes. ¡Siempre es bueno contar con una segunda opinión!”(9).

A la pregunta “¿son tus cálculos lo suficientemente confiables para ser usados ante un tribunal?”, el chatbot respondió: “Cuando se trata de asuntos legales, cualquier cálculo o dato debe cumplir con estándares estrictos. Puedo proporcionar información precisa, pero siempre debe ser verificada por expertos y acompañada de evaluaciones profesionales antes de ser utilizada en un tribunal”(10). En suma, Copilot “declaró” que depende de la supervisión y análisis humanos.

El juez recordó, hace tiempo, que la ley del estado de Nueva York aplica el estándar del precedente Frye al testimonio de expertos y la prueba científica(11). Al estudiar y revisar los casos y las reglas de práctica

judicial en los Estados Unidos, sostuvo que debido a la rápida evolución de la inteligencia artificial y su inherente falta de confiabilidad, antes de presentar pruebas generadas por sistemas o productos de inteligencia artificial, los abogados tienen la obligación de revelar el uso de la inteligencia artificial y las pruebas que se pretendan presentar deberán cumplir con el estándar de Frye. Para ello, se llevará a cabo una audiencia de admisión de tales pruebas (antes del juicio o al momento de producirlas) donde se presentará testimonio de expertos, se explicarán los procesos de obtención de las pruebas, las fórmulas matemáticas si las hubiera, y se hará referencia a las publicaciones científicas(12).

IV. Conclusión

Las respuestas obtenidas por el juez demuestran que los desarrolladores de Copilot en Microsoft reconocen la necesidad de su supervisión por un operador (persona humana) para comprobar la exactitud de los datos suministrados (input) como así también el resultado obtenido (output). La calidad de las respuestas que los sistemas de inteligencia artificial brindan depende de los datos (input) que se les proporciona y de cómo los estructura. La inteligencia artificial generativa busca información en los macrodatos (big data) y confecciona una respuesta en el estilo que se le ha solicitado. No desarrolla conceptos ni análisis nuevos. Repite lo que encuentra y lo presenta en un formato atractivo. Reordena contenidos preexistentes y contribuye a consolidarlos, sin examinar si contienen errores o prejuicios.

El tribunal federal de apelaciones del distrito de Columbia (más conocido como D.C. Circuit) sostuvo, en el precedente Frye, que el testimonio de expertos debe basarse en métodos científicos que estén suficientemente establecidos y aceptados. La fecha de esa sentencia es el 3 de diciembre de 1923.

El caso versó sobre la admisibilidad de la prueba de engaño de la presión arterial sistólica. Durante el juicio, un experto argumentó que los experimentos científicos habían demostrado que el miedo, la rabia, y el dolor siempre producen un aumento de la presión arterial sistólica; mientras que el engaño, el ocultamiento de los hechos o la culpabilidad de un delito, acompañados del miedo a ser descubierto cuando el imputado está bajo examen, eleva la presión arterial sistólica en una curva que se corresponde con el intento de controlar el miedo y engañar al examinador. La prueba no fue admitida por carecer de aceptación científica.

Un siglo más tarde, la prueba producida con asistencia de la inteligencia artificial deberá basarse también en métodos científicos suficientemente aceptados y establecidos para ser admisible ante el tribunal del condado de Saratoga, NY. La habilidad de la inteligencia artificial para producir textos sintáctica y semánticamente coherentes (o de producir cálculos en el asunto de Weber) no es garantía de fiabilidad.

(*) Profesora de Derecho Internacional Público en la Facultad de Derecho de la UCA. Abogada (UCA, 2005). Master of Laws in International Law (The Fletcher School of Law and Diplomacy, 2009) como becaria Fulbright 2008-2009.

VOCES: ABOGADO - DEMANDA - PROCESO JUDICIAL - PODER JUDICIAL - NOTIFICACIÓN - INFORMÁTICA - TECNOLOGÍA - CORTE SUPREMA DE LA NACIÓN - DOMICILIO - DEFENSA EN JUICIO - EXPEDIENTE JUDICIAL - EJERCICIO PROFESIONAL - JUECES - PRUEBA - VALORACIÓN DE LA PRUEBA - SENTENCIA - PROCESO ORDINARIO - RECURSOS PROCESALES - JURISPRUDENCIA - CONSTITUCIÓN NACIONAL - EJERCICIO PROFESIONAL - COMUNICACIONES ELECTRÓNICAS - INTERNET - INTELIGENCIA ARTIFICIAL - DERECHOS Y GARANTÍAS CONSTITUCIONALES - ACCESO A LA JUSTICIA - PERSONAS FÍSICAS - TRIBUNALES EXTRANJEROS

(1) In the Matter of the Accounting by Susan F. Weber as the Trustee of the Michael S. Weber Trust for the benefit of Owen K. Weber under Article Seventh of the Last Will and Testament of Michael S. Weber, Decedent, File No. 1845-4/B (Decided on October 10, 2024), Surrogate's Court, Saratoga County (NY), published by the New York State Law Reporting Bureau pursuant to Judiciary Law §431.

(2) Para un perfil del Sr. Ranson, véase: <https://charleswranston.com/> (último acceso el 16 de noviembre de 2024).

(3) Debe recordarse que el experto calculó daños desde 2004, aunque corresponda calcularlos desde el 25 de enero de 2008.

(4) "A. M. R." significa "Account Management Resources".

(5) Matter of Weber (2024 NY Slip Op 24258), p. 29. Traducción de la autora. "The use of artificial intelligence is a rapidly growing reality across many industries".

(6) *Ibidem*. Traducción de la autora. "The mere fact that artificial intelligence has played a role, which continues to expand in our everyday lives, does not make the results generated by artificial intelligence admissible in Court".

(7) Unified Court System (UCS) issued computer.

(8) Supra 5, p. 27. Traducción de la autora. "I am to be accurate within the data I've been trained on and the information I can find for you. That said, my accuracy is only as good as my sources so for critical matters, it's always wise to verify".

(9) Ibídem. Traducción de la autora. "[Y]ou bet. When it comes to providing information and engaging in conversation, I do my best to be as reliable as possible. However, I am also programmed to advise checking with experts for critical issues. Always good to have a second opinion!".

(10) Ibídem, p. 28. Traducción de la autora. "[W]hen it comes to legal matters, any calculations or data need to meet strict standards. I can provide accurate info, but it should always be verified by experts and accompanied by professional evaluations before being used in court".

(11) Frye vs. United States, 293 F 1013 [D.C. Cir. 1923].

(12) Para un ejemplo de ello, véase People vs. Wakefield, 38 NY3d 367 [2022]. En este caso penal, se permitió el empleo de un programa de software con asistencia de inteligencia artificial para realizar un análisis de ADN.

© Copyright: El Derecho
