LITORALPRESS

Fecha: 22-04-2024
Medio: El Mercurio
Supl. : El Mercurio - Cuerpo A
Tipo: Noticia general
Título: Deepfakes y robollamadas: la amenaza de la IA se toma las elecciones estadounidenses

Pág. : 5
Cm2: 772,2

Tiraje: 126.654
Lectoría: 320.543
Favorabilidad: No Definida

Empresas han prometido combatirlas y se han multiplicado las leyes al respecto

Deepfakes y robollamadas: la amenaza de la IA se toma las elecciones estadounidenses

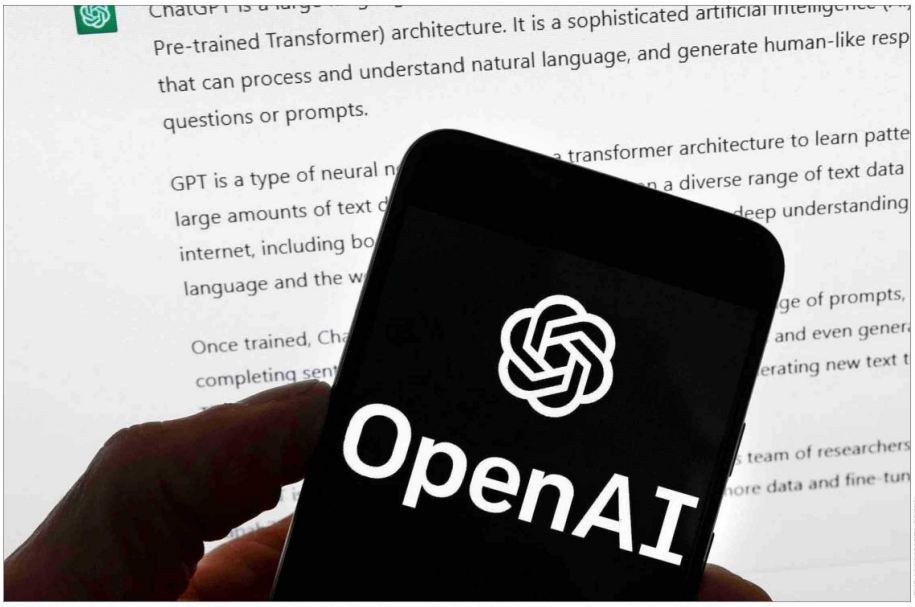
La masividad que ha tomado esta tecnología en el último tiempo y su facilidad de uso preocupan a las autoridades.

NICOLÁS GARCÍA DE VAL

Si 2023 fue el "año de la Inteligencia Artificial (IA)" por la explosiva entrada de esta tecnología al uso cotidiano de millones de personas, con herramientas como ChatGPT, este podría convertirse en el "año de la desinformación con IA". Países alrededor de todo el mundo han enfrentado una propagación de "noticias falsas" en sus procesos electorales en 2024 y la situación podría volverse crítica en Estados Unidos, donde el Presidente Joe Biden y su predecesor, Donald Trump, enfrentan la que puede ser una competencia voto a voto por el liderazgo de la principal potencia mundial.

La primera señal de la gravedad del problema en el país llegó en enero, cuando se alertó que los votantes en la primaria de New Hampshire estaban siendo objeto de una campaña de desinformación de robollamadas generadas con IA. En el telefonazo se escuchaba, supuestamente, a Biden diciéndoles que se quedaran en casa. Todavía no hay claridad sobre si esta iniciativa de desinformación tuvo un impacto real y llevó a los votantes a no ir a las urnas, pero el gobierno local está investigando lo ocurrido y ya hubo demandas contra Lingo Telecom y Life Corporation, las dos empresas texanas que, se cree, están vinculadas a lo ocurrido.

Las robollamadas no son la única amenaza, pues los ojos de todos están puestos en las *"deepfakes"*, como se conoce a los videos falsos hiperrealistas creados con IA, que se han convertido en la última amenaza tecnológica en lo que a noticias falsas se refiere. Ya se han visto casos puntuales de videos *deepfake* en esta campaña y aunque todavía no ha habido uno masivo a nivel nacional, los casos que han ocurrido en elecciones en el extranjero tienen a las autoridades en alerta.



THE ASSOCIATED PRESS

LAS GRANDES compañías de IA están en el centro de la discusión por el potencial mal uso de sus herramientas durante la campaña.

En octubre de 2023 en Eslovaquia, por ejemplo, una imitación de un líder opositor podría haber ayudado a ganar a un candidato pro ruso. También se han visto casos en Pakistán —donde una *deepfake* de un candidato llamó a las personas a boicotear la elección de febrero de este año— y en Bangladesh.

Esto se ve potenciado por la enorme dificultad que tienen las personas en darse cuenta cuando están viendo contenido generado por IA. Un estudio reciente de la Northeastern University, por ejemplo, mostró que 52% de las personas dice no poder distinguir entre videos e imágenes generadas por un humano y aquellas que son hechas por IA. Y el problema solo se volverá más grave a medida que aumente el desarrollo de esta tecnología.

"La desinformación generada por IA es una grave amenaza este año porque permite generar rápida y fácilmente videos y audios falsos. No hace falta ser un experto técnico para utilizar la IA generativa, cualquier puede usarla", dijo a "El Mercurio" Darrell West, miembro de la Brookings Institution y exdirector del Centro de Innovación Tecnológica de la organización.

Ni siquiera es necesario que haya actores maliciosos de por medio para que esta tecnología sea una amenaza electoral. Las aplicaciones de IA son capaces de crear textos bien redactados y útiles, pero también se suelen confundir y escribir textos que son "confiadamente incorrectos". Es lo que se conoce como "verdaderidad" (*"truthiness"* en inglés), concepto que se usa para describir información que se percibe como correcta, sin necesariamente serlo. Un estudio de AI Democracy Projects y Proof News mostró que herramientas populares de IA —incluyendo Gemini de Google, ChatGPT y el Llama 2 de Meta— "tuvieron un mal desempeño en precisión" cuando les hicieron ciertas preguntas sobre las elecciones.

Acciones necesarias

En ese escenario, todos los ojos están puestos sobre las grandes tecnológicas que desarrollan aplicaciones de este tipo y que en los últimos años han integrado la IA a distintas partes de sus operaciones y servicios a los usuarios.

En febrero, un grupo de 20 de estas compañías anunció que habían acordado trabajar juntas para prevenir que contenido creado por IA interfiriera en las elecciones a nivel mundial. Entre las firmantes están OpenAI (dueños de ChatGPT), Microsoft, Adobe, Meta, Tiktok y X. El acuerdo incluye compromisos para colaborar en el desarrollo de herramientas para detectar imágenes, audios y videos generados por IA que puedan provocar desinformación, y tomar medidas para evitar que estos contenidos se generen.

Sin embargo, esto no es suficiente.

"No basta simplemente con poner trabas para que se generen textos con IA sobre elecciones o figuras políticas. Hemos visto una y otra vez lo fácil que es saltarse estas trabas. Tampoco basta con decir simplemente en sus términos de servicio que el sistema no debe utilizarse para interferir en las elecciones. Estas empresas deben contar con mecanismos sólidos para monitorear y detener los usos de la IA que ellas mismas establecen como usos prohibidos. Y, por último, necesitamos centrarnos más en los mecanismos de etiquetado y divulgación del contenido generado por IA", dijo a este diario Merve Hickock, presidente del Centro para Política Digital y de IA y profesora de la Escuela de Información de la Universidad de Michigan.

Otros expertos plantean que las grandes tecnológicas no solo no han hecho suficiente, sino que además tienen un incentivo para permitir la desinformación. "Empresas como Meta ganan dinero cuando sus usuarios participan activamente, así pueden ser vigilados más de cerca y enviarles anuncios con mayor frecuencia. Cosas como la desinformación aumentan la interacción. Por eso se sienten incentivadas a prometer mucho, pero hacer poco", manifestó a "El Mercurio" Bruce Schneier, un renombrado especialista en tecnología y seguridad de la Universidad de Harvard.

Ante la aparente incapacidad de las grandes empresas tecnológicas de frenar la difusión de contenido malicioso generado con IA, los legisladores estadounidenses han intentado hacer algo al respecto.

Congresistas locales de 40 estados de EE.UU. han presentado proyectos de ley que buscan regular el uso de *deepfakes* durante las elecciones, 21 de ellos han sido votados por al menos una Cámara, y diez han aprobado normas al respecto, según la ONG Public Citizen.

Muchas de estas leyes se enfocan en la transparencia, y ordenan a las campañas y candidatos que revelen si utilizan videos, audios o imágenes generadas con IA. Otras prohíben el uso de *deepfakes* 60 o 90 días antes de las elecciones.

Una de estas fue firmada en Wisconsin a fines de marzo por el gobernador Tony Evers. La norma exige a los grupos afiliados con campañas políticas que incluyan avisos de contenidos generados por IA. De no hacerlo arriesgan una multa de US$ 1.000 por cada violación a la ley.

Este tipo de reglas ayudan, pero los expertos plantean que van demasiado lento. "Las regulaciones gubernamentales para la IA aún están en sus primeras etapas, pero la tecnología está evolucionando rápidamente. Nos acercamos a un punto en el que las personas tendrán dificultades para distinguir la verdad de la ficción en imágenes, videos y audio creados por IA. Las regulaciones efectivas requerirán la colaboración entre los gobiernos, las empresas de inteligencia artificial que desarrollan tecnología *deepfake* y las plataformas de redes sociales donde se difunde la desinformación. Los rápidos cambios dificultan la creación de regulaciones que sean capaces de combatir la desinformación antes de las elecciones estadounidenses de 2024", manifestó Sarah Studer de True Media, una organización que estudia *deepfakes* y desarrolló una herramienta para detectarlas.