

Inteligencia artificial, elecciones, medios y desinformación

La desinformación creada con inteligencia artificial generativa ya **es una realidad que intercede en resultados electorales**. Este tipo de desinformación, que afectó a uno de los principales candidatos de las elecciones de 2023 en Eslovaquia, representa un grave riesgo para los sistemas democráticos. Organizaciones dedicadas a la verificación de datos destacan la necesidad de realizar esfuerzos constantes para combatir estas amenazas, así como de establecer estándares claros en el periodismo y las plataformas digitales para **proteger la integridad de los procesos democráticos**.

CARLOS HERNÁNDEZ-ECHEVARRÍA

Uno de los peores escenarios posibles empieza así: en la mañana de la jornada de reflexión, un teléfono vibra. Es una nota de voz. A la hora del aperitivo, el mensaje está ya en 100.000 teléfonos; y después de la siesta, en medio millón. En los grupos de WhatsApp de los amigos del colegio, de la familia, de los padres de la guardería... Una grabación en la que uno de los candidatos de los principales partidos le dice a alguien que no se preocupe, que las elecciones están amañadas. Es su voz, es la misma voz que todo el mundo ha escuchado millones de veces, pero no es su voz. Está generada por inteligencia artificial (IA).

El mensaje vuela. A media tarde, tres millones de españoles ya han escuchado la grabación y la discusión ha saltado a las redes sociales. Unos dicen que la voz está generada por IA, otros que no y otros que, incluso si no lo ha dicho, probablemente es cierto. En muchas webs ya se ha publicado la “controvertida grabación” sin aclarar si es cierta. Los medios de comunicación están haciendo averiguaciones, pero no es fácil encontrar expertos que puedan afirmar categóricamente una cosa o la otra. El partido político del candidato saca un comunicado antes de la hora de la cena diciendo que no es real; el partido rival

Carlos Hernández-Echevarría es director adjunto de *Maldita.es*

saca otro que no dice que lo sea, aunque pide una investigación ante unas declaraciones “gravísimas de ser ciertas”.

A la mañana siguiente, seis millones de españoles van a votar habiendo escuchado el audio. Tres millones saben que el candidato lo niega y quizás unos 400.000 han leído artículos en profundidad en los que se explica por qué ya está claro que el mensaje partió de una IA y que no es auténtico. A las nueve, cierran las urnas y todas las opciones son malas: tal vez cientos de miles de españoles, si no millones, no reconocen el resultado, porque aún creen que la grabación es auténtica. O tal vez otros cientos de miles de españoles, si no millones, no reconocen el resultado, pues creen que “les han robado las elecciones” con una grabación falsa.

Aunque parezca el inicio de una ficción distópica, no lo es en absoluto. Ya ha pasado: hace nueve meses a uno de los dos principales candidatos en Eslovaquia le sucedió exactamente eso. Durante los últimos días de campaña, alguien hizo circular un audio viral generado con IA en el que decía que había amañado las elecciones¹ y también otro en el

que “confesaba” que iba a subir el precio de la cerveza. Y perdió. ¿Fue por eso? Nadie lo sabe a ciencia cierta, si bien se estima que en un país de 5,5 millones de habitantes el audio le llegó a cientos de miles de personas².

Existe una clara voluntad por parte de algunos de usar mentiras para atacar el sistema electoral

Los verificadores independientes como *Maldita.es* estamos intentando estar preparados para este escenario, trabajando con herramientas tecnológicas mejores (muchas de ellas también usan IA) para detectar los bulos antes y tardar menos en desmentirlos; también para conseguir que nuestros desmentidos lleguen más rápido a la gente, preferentemente antes que la desinformación. Tenemos equipos especializados y los reforzamos en los días previos a las elecciones, y contamos además con el apoyo de expertos externos que nos ayudan con sus superpoderes³ y a los que les pedimos que estén

1 Devine, C., O’Sullivan, D. y Lyngaas, S. (1 de febrero de 2024). “Una grabación falsa de un candidato diciendo que había manipulado las elecciones se volvió viral. Los expertos dicen que es solo el comienzo”. CNN. <https://edition.cnn.com/2024/02/01/politics/election-deepfake-threats-invs/index.html>

2 Köváry Sólymos, K. (8 de noviembre de 2023). “Las elecciones eslovacas son blanco de un engaño profundo pro-Kremlin”. *ICJK*. <https://vsquare.org/slovak-election-targeted-by-pro-kremlin-deepfake-hoax/>

3 *Maldita.es*. Comunidad de Maldita. “Tienes un superpoder para luchar contra los bulos, ¡compártelo!”. <https://comunidad.maldita.es/u/comparte-superpoderes>

especialmente disponibles para atendernos en esos días. No obstante, hay que reconocer que una cosa así puede pasar y que podría hacer un daño inmenso.

Idealmente, todos deberíamos saber ya que las cosas que nos llegan al teléfono no tienen por qué ser ciertas, pero a la vez el mensaje tampoco puede ser que no hay que fiarse de nada ni de nadie. Hay todo un ecosistema informativo que tiene que responder a ese reto y no solo durante las dos semanas de campaña electoral. Las elecciones son una prioridad para los desinformadores, que ahora cuentan con herramientas que facilitan que su trabajo llegue a más gente con menos esfuerzo, así que aquellos que luchamos contra la desinformación también tenemos que mejorar nuestras capacidades ante una amenaza que crea un riesgo real para el sistema democrático y para las personas que viven en él.

La desinformación sobre elecciones es particularmente peligrosa

El papel de la inteligencia artificial generativa es nuevo, pero la desinformación electoral viene de mucho antes. Y probablemente, la más peligrosa no es la desinformación sobre los candidatos y los partidos, sino la que tiene como objetivo el proceso electoral en sí. Desde hace años, hemos visto bulos que intentan socavar la confianza de los ciudadanos en las elecciones usando métodos que no tenían nada de sofisticados: vídeos fuera de contexto⁴, imágenes alteradas⁵, falsedades sobre el procedimiento de conteo⁶... Hemos aclarado el papel de Correos o de Indra en el proceso electoral cuando se acusaba a Rajoy⁷ de usarlos para perpetrar un fraude y cuando esa acusación se desplazó a Pedro Sánchez⁸.

Existe una clara voluntad por parte de algunos de usar mentiras para atacar el

4 *Maldito Bulo* (23 de junio de 2023). "No, este vídeo en el que una mujer entrega 'un tocho de papeletas' en una oficina de Correos no es de León ni actual: es Melilla en 2015". *Maldita.es*. <https://maldita.es/malditobulo/20230623/oficinas-correos-leon-tocho-papeletas-pucherazo/>

5 *Maldito Bulo* (23 de julio de 2023). "No, esta foto de Ángela Rodríguez 'Pam' con una papeleta con el nombre de Irene Montero escrito a mano no es real". *Maldita.es*. <https://maldita.es/malditobulo/20230723/angela-rodriguez-pam-papeleta-montero/>

6 *Maldito Bulo* (28 de julio de 2023). "La ley no dice que hay que contar de nuevo una a una las papeletas y que estas no se eliminen: se revisan las actas". *Maldita.es*. <https://maldita.es/malditobulo/20230728/contar-votos-uno-eliminacion-escrutinio-pucherazo/>

7 *Maldito Bulo* (26 de abril de 2017). "No, Indra no hace el recuento de votos en las elecciones en España". *Maldita.es*. <https://maldita.es/malditobulo/20170426/no-indra-no-hace-el-recuento-de-votos-en-las-elecciones-en-espana/>

8 *Maldito Bulo* (1 de marzo de 2023). "Las desinformaciones sobre Indra y las elecciones en España". *Maldita.es*. <https://maldita.es/malditobulo/20230301/indra-elecciones-recuento-votos/>

sistema electoral. Tal vez su objetivo sea deslegitimar los resultados y a las instituciones que salen de ellos, o quizás solo buscan transmitir la idea de que tu voto no importa porque esos resultados ya están decididos de antemano. Lo que está claro es que, cuando la desinformación sobre el proceso electoral convence a millones de que el resultado es ilegítimo, esas personas pueden recurrir a la violencia para dar la vuelta a lo que consideran una maniobra antidemocrática: lo hemos visto en el Capitolio de EE. UU. en 2021 y en el Congreso de Brasil al año siguiente. El riesgo es real.

En la campaña para las elecciones al Parlamento Europeo, ya detectamos desinformación contra el proceso electoral en varios países europeos. Verificadores independientes en Francia, Alemania y España desmintieron contenidos sobre supuestos “pucherazos”⁹, papeletas electorales manipuladas¹⁰ y otras falsedades varias semanas antes de las elecciones. A veces son bulos que vuelven periódicamente, asociados una y otra vez al mismo vídeo sacado de contexto o a la misma imagen manipulada.

Es difícil no pensar que parte de estas campañas vienen de actores organiza-

dos que no están interesados en promover un bulo concreto, sino una narrativa general de que las democracias europeas no son reales y que los votantes no eligen realmente a quien toma las decisiones. Que las democracias no son, dicho de otro modo, tan diferentes de los regímenes autoritarios. Es una idea que se repite una y otra vez, y no solo durante las dos semanas que dura la campaña electoral legal. Cabe también preguntarse quién puede tener interés en fomentar estas narrativas falsas con tanta insistencia.

Las elecciones son mucho más que la campaña

En la pasada campaña para las elecciones europeas, hemos visto multitud de iniciativas contra la desinformación por parte de las instituciones de la Unión Europea (UE), los Estados miembros, las grandes plataformas digitales..., pero esas actuaciones circunscritas al periodo electoral legal no van a marcar una enorme diferencia. Aunque muchas personas deciden si van a participar o a quién votar durante los últimos días antes de las elecciones, las opiniones generales sobre la marcha de la sociedad y la calidad del sistema democrático se crean durante un

⁹ Nicolaus, K. (10 de mayo de 2024). “Groß Schirma: No fueron los ‘viejos partidos’, sino los expertos jurídicos quienes declararon inválidas las elecciones a la alcaldía”. *Correctiv*. <https://correctiv.org/faktencheck/2024/05/10/grossschirma-nicht-altparteien-sondern-juristische-fachleute-erklaerten-die-buergermeisterwahl-fuer-ungueltig/>

¹⁰ Lehn, J. (15 de mayo de 2024). “Un agujero en la papeleta electoral de la UE ayuda a los ciegos, el voto sigue siendo válido”. *AFP Faktencheck*. <https://faktencheck.afp.com/doc.afp.com.34RH496>

tiempo mucho más largo. Por ello, la pelea contra la desinformación que afecta a las elecciones hay que hacerla todos los días, no solo cuando están cerca.

Las campañas de desinformación más efectivas son continuadas y no saben de periodos electorales fijos: son aquellas que, como decíamos, empujan una narrativa general a través del tiempo para construir una percepción, una idea general más poderosa que un bulo concreto en un momento concreto. El ecosistema informativo tiene, por tanto, que estar preparado para ese tipo de reto: los medios de comunicación, las plataformas digitales, las iniciativas de educación mediática... todos esos espacios tienen que acoger un esfuerzo contra la desinformación bien planificado, efectivo y estable en el tiempo. Solo con una campaña institucional a 15 días de las elecciones no se puede hacer mucho.

En el caso de las últimas elecciones, la UE estrenó su flamante nuevo marco regulatorio para las grandes plataformas digitales y buscadores, la Ley de Servicios Digitales¹¹. Bajo esta regulación, es-

tas grandes empresas están obligadas a tener medidas de mitigación de riesgos efectivas contra males sistémicos del mundo *online* como la desinformación¹², prestando una atención especial a prevenir la manipulación de las elecciones. Y, sin embargo, ¿las tienen?

Las campañas de desinformación más efectivas son continuadas y no saben de periodos electorales fijos

Si volvemos al ejemplo de los vídeos generados por inteligencia artificial, incluso las empresas tecnológicas más importantes del mundo no están seguras de su propia capacidad de detectar¹³ si una grabación está hecha por una IA o no. No obstante, sus limitaciones a la hora de abordar el contenido falso y peligroso van más allá de ese caso concreto. Muchas de ellas, como YouTube, obstaculizan el trabajo de los verificadores in-

11 Parlamento Europeo y Consejo de la Unión Europea (2022). Reglamento (UE) 2022/2065, de 19 de octubre, relativo al mercado único de los servicios digitales y por el que se modifica la Directiva 2000/31/CE. Diario Oficial de la Unión Europea, L 277, 1-102. <https://eur-lex.europa.eu/eli/reg/2022/2065/oj>

12 *Maldita Tecnología* (2 de enero de 2024). "Todo sobre la Ley de Servicios Digitales (DSA): qué es, cómo nos afecta y cómo se está implementando". *Maldita.es*. <https://maldita.es/malditatecnologia/20240102/ley-servicios-digitales-dsa-como-se-esta-implementando/>

13 Pugalia, A. et al. (2024). "The Race to Detect AI-Generated Content and Tackle Harms". *Tech Policy Press*. <https://www.techpolicy.press/the-race-to-detect-ai-generated-content-and-tackle-harms/>

dependientes¹⁴ mientras borran millones de vídeos que consideran desinformación sin explicar nunca cómo se toman esas decisiones ni por qué alguien (¿una persona?, ¿un algoritmo?) ha llegado a la conclusión de que una determinada afirmación es falsa.

Estas desventajas estructurales no se pueden arreglar en 15 días de campaña, incluso si la intención de hacerlo fuera genuina. Lo mismo pasa con otra parte clave del ecosistema informativo: los periodistas y los medios de comunicación. Las carencias del sector son una oportunidad para las campañas de desinformación antes, durante y después de cada elección. El hecho de que en España no exista ni un mínimo consenso sobre qué es un medio de comunicación y qué no lo es representa quizás el ejemplo más relevante de estas debilidades.

Hay multitud de webs con aspecto de medio de comunicación que se dedican a publicar mentiras de forma constante y que se aprovechan de esa apariencia para resultar más creíbles. Y esa estructura va a estar siempre ahí para difundir desinformación: durante los procesos electorales, antes y después. Hasta hoy no hemos sido capaces de establecer qué es lo mínimo que se debería exigir para poder llamar a algo “medio de comuni-

cación”. Unos estándares que valgan para medios grandes y pequeños, con una línea editorial o con otra, pero que supongan una separación efectiva con respecto a webs y canales que deben poder publicar lo que quieran, si bien no deberían tener el reconocimiento público de ser llamados “medios de comunicación”.

Las carencias del sector son una oportunidad para las campañas de desinformación antes, durante y después de cada elección

La solución no puede ser que el Gobierno de turno venga a decir quién es un medio y quién no, es una reflexión que tiene que empezar entre periodistas, medios, académicos, verificadores... Quizás esos estándares para medios deben empezar por la transparencia: quién manda, quién paga, quién toma las decisiones editoriales; pero también por unos estándares éticos sólidos, incluyendo una política de rectificación clara, transparente y con consecuencias si no se lleva a cabo. Si de cara a las elecciones del futuro tenemos una distinción clara

14 Associated Press (AP) (13 de enero de 2022). “Los verificadores de hechos dicen que YouTube permite que su plataforma sea ‘un arma’”. <https://apnews.com/article/technology-business-arts-and-entertainment-misinformation-susan-wojicki-ed9fc345e037e13b87955d7a398233bf>

y creíble entre medios de comunicación y cosas que no lo son, podremos pedir a los ciudadanos que tomen decisiones informadas al respecto.

Esto puede tomar la forma de un organismo independiente y profesional que no dependa de las instituciones y que vigile si los medios que dicen aplicar esos estándares lo hacen en realidad. Si rectifican ante sus audiencias cuando se equivocan y si son transparentes en cuanto a su propiedad y toma de decisiones. Y si no lo hacen, ese órgano debería retirarles esa distinción, ese sello de calidad, para que las audiencias y los anunciantes sepan a qué atenerse cuando vean sus contenidos. No se trata nunca de impedir a nadie publicar, solo de dar una distinción a quien sí cumple determinados mínimos.

Los riesgos de la IA en la desinformación van mucho más allá de los ‘deepfakes’

En los próximos años, todo indica que el ecosistema informativo estará cada vez más sometido a la mediación de la inteligencia artificial generativa, como en los últimos años lo ha estado a la de los motores de búsqueda y recomendación¹⁵. Esta influencia en los modos de adquisición de información jugará un papel muy relevante en la creación de

opiniones y, por tanto, en las elecciones, según la gente vaya buscando las respuestas a sus preguntas más y más en sistemas que usen IA generativa. Merece la pena preguntarse de dónde saldrán esas respuestas y con qué información se elaborarán.

Actualmente, los sistemas más avanzados de IA generativa que están a disposición del público para responder preguntas son los de OpenAI y Google Gemini. Sus grandes modelos de lenguaje se entrenan principalmente sobre contenido público, es decir, están aprendiendo lo que es el mundo a través del contenido *online* que está disponible, no del contenido *online* que es factual o de calidad. De hecho, estas herramientas, que van a tener una importancia crucial en la creación de opinión, van a hacer un juicio sobre la realidad que tiene que ver más con qué versión es más popular o aparece más en internet que con qué versión es real y ha sido verificada como tal.

Pensemos en esto desde la perspectiva de los medios de comunicación. El periodismo es un enorme creador de datos de alto valor, comprobados, con una supervisión humana robusta y recogidos por una metodología clara. Es, en muchos aspectos, el estándar de oro y lo

¹⁵ González Alba, J. A. (2022). “La adaptación de los profesionales a la transformación digital de los medios: los nuevos perfiles que se incorporan a las redacciones”. *Cuadernos de Periodistas*, nº 43. <https://www.cuadernosdeperiodistas.com/la-adaptacion-de-los-profesionales-a-la-transformacion-digital-de-los-medios-los-nuevos-perfiles-que-se-incorporan-a-las-redacciones/>

que les falta a esos sistemas de IA para poder ofrecer resultados de mayor calidad a sus usuarios. Sin embargo, y dado que los medios (por buenas razones) protegen sus datos cada vez más detrás de muros de pago, cabe preocuparse por la calidad de lo que queda detrás y que servirá para entrenar a los modelos del futuro. La desinformación es casi siempre gratuita y accesible.

Algunos grandes grupos de medios están llegando a acuerdos con las compañías punteras de la IA generativa para usar su contenido para entrenar sus modelos¹⁶; pero, aun así, esas colaboraciones representan una pequeñísima parte de la producción de los medios. Del lado de las organizaciones independientes de verificación como *Maldita.es*, se está explorando también cómo el trabajo de los *fact-checkers* podría ayudar a que determinados modelos ofrecieran menos desinformación, así como para usar la IA generativa en la identificación de pa-

trones y narrativas que ya afectan hoy a las elecciones.

La desinformación es casi siempre gratuita y accesible

Las muchas elecciones de este 2024, en particular las europeas y las presidenciales en EE. UU., representan una oportunidad para quienes quieren usar las nuevas capacidades tecnológicas de la IA para manipular a través de la desinformación, sobre todo ahora que esas herramientas son relativamente nuevas y los votantes están menos alerta ante ellas. Del lado del periodismo, nuestra responsabilidad es entenderlas, usarlas en sentido contrario y prevenir a la ciudadanía de sus posibilidades; pero también trabajar a favor de un ecosistema informativo en el que la gente tenga referencias más claras de confianza dentro de los medios de comunicación para que no se la cuelen. ■

¹⁶ Robertson, K. (22 de mayo de 2024). "OpenAI llega a un acuerdo para licenciar el contenido de News Corp". *The New York Times*. <https://www.nytimes.com/2024/05/22/business/media/openai-news-corp-content-deal.html>