

El discurso de los periodistas sobre el impacto de la inteligencia artificial generativa en la desinformación¹

Simón Peña-Fernández², Urko Peña-Alonso³ y Mainer Eizmendi-Iraola⁴

Recibido: 10 de mayo de 2023 / Aceptado: 6 de agosto de 2023

Resumen. La inteligencia artificial generativa ha irrumpido con fuerza en el curso de la transformación digital de los medios de comunicación como una herramienta tecnológica de naturaleza dual, que ofrece nuevas posibilidades para combatir la desinformación, pero también para generarla. Esta investigación analiza la opinión de 118 profesionales de la información sobre el impacto de la IA, en particular en lo referido a la desinformación, a partir de 778 mensajes publicados en Twitter. Los resultados señalan que los periodistas españoles expresan una opinión mayoritariamente cauta que huye de posturas tecnocéntricas y en la que prevalece la preocupación por el modo en el que la IA puede afectar a la ciudadanía. Los principales peligros percibidos son el incremento de los desórdenes informativos y la desinformación, los sesgos en los resultados obtenidos, la inexactitud de los contenidos, la facilidad para cometer fraude o elaborar deepfakes, o su potencial uso con fines maliciosos.

Palabras clave: IA, inteligencia artificial generativa; desinformación; medios de comunicación; periodistas

[en] The discourse of journalists on the impact of generative artificial intelligence on disinformation

Abstract. Generative artificial intelligence has broken into the course of the digital transformation of the media as a technological tool of a dual nature, offering new possibilities to combat disinformation, but also to generate it. This research analyzes the opinion of 118 information professionals on the impact of AI, particularly with regard to misinformation, based on 778 messages published on Twitter. The results indicate that Spanish journalists express a mostly cautious opinion that flees from technocentric positions and in which concern about the way in which AI can affect citizens prevails. The main perceived dangers are the increase in information disorders and misinformation, bias in the results obtained, the inaccuracy of the content, the ease of committing fraud or developing deepfakes, or their potential use for malicious purposes.

Keywords: AI; generative artificial intelligence; disinformation; media; journalists

Sumario. 1. Introducción 2. Metodología 3. Resultados 3.1. Difusión 3.2. Potencialidades y riesgos 3.3. El impacto de la IA en la desinformación 4. Conclusiones y discusión 5. Referencias bibliográficas

Cómo citar: Peña-Fernández, S., Peña-Alonso, U. & Eizmendi-Iraola, M. (2023). El discurso de los periodistas sobre el impacto de la inteligencia artificial generativa en la desinformación. *Estudios sobre el Mensaje Periodístico* 29 (4), 833-841. <https://dx.doi.org/10.5209/esmp.88673>

1. Introducción

La desinformación se ha convertido en una de las principales preocupaciones en el ámbito de la comunicación durante la última década (Posetti & Matthews, 2018; Waisbord, 2018; Cea & Palomo, 2021). A su difusión ha contribuido el incremento del número de actores que pueden distribuir información en un sistema de medios que ha evolucionado desde un modelo altamente jerarquizado con un número muy limitado de fuentes que se dirigían a una audiencia eminentemente pasiva, a un modelo híbrido en

el que múltiples actores con intereses diferenciados se dirigen a unas audiencias activas y empoderadas (Holton & Belair Gagnon, 2018; Masip, Ruiz-Caballero & Suau, 2019).

En el complejo ecosistema informativo digital, las audiencias se enfrentan a la paradoja de que, con el incremento del volumen de información disponible, aumentan también las vulnerabilidades y los desórdenes informativos (Wardle & Derakhshan, 2017), la polarización y la exposición a contenidos falsos (Marwick y&Lewis, 2017), y desciende el pluralismo (Sunstein, 2017).

¹ Este trabajo forma parte del proyecto de investigación TED2021-130810B-C22 financiado por Agencia Estatal de Investigación (AEI) del Ministerio de Ciencia e Innovación (Micin) de España 10.13039/501100011033 y por la Comisión Europea Next Generation EU/PRTR. También forma parte de la producción científica del Grupo de Investigación Consolidado del Sistema Universitario Vasco Gureiker (IT1496-22).

² Universidad del País Vasco (España)
E-mail: simon.pena@ehu.es

³ Universidad del País Vasco (España)
E-mail: urko.pena@ehu.es

⁴ Universidad del País Vasco (España)
E-mail: mainer.eizmendi@ehu.es

El desarrollo tecnológico en el ámbito de la comunicación ha mostrado su carácter multiplicador, tanto en las posibilidades de generar desinformación, como de combatirla mediante herramientas automatizadas (Flores-Vivar, 2019) o a través de nuevos roles profesionales como los *fact-checkers* (Clayton et al., 2020).

En este contexto, la irrupción de las nuevas herramientas de inteligencia artificial generativa ha reavivado el debate sobre la naturaleza dual de los desarrollos tecnológicos (Brennen, Howard & Nielsen, 2018) o, lo que es lo mismo, el equilibrio entre los beneficios que suponen y las nuevas vulnerabilidades que crean.

A la aplicación de la inteligencia artificial se le atribuyen potencialmente beneficios en todos los procesos industriales, incluidos los comunicativos, pero también en el ámbito social (Comisión Europea, 2020). Así, su desarrollo ha favorecido la aparición de herramientas automatizadas que permiten combatir la desinformación en las redes sociales (Zhuk et al., 2018; Altunbey & Alatas, 2020). Gracias a ellas, los medios y las organizaciones de *fact-checking* pueden ampliar el rango de sus verificaciones, más allá de la comprobación manual o el simple etiquetado de mensajes sospechosos, en particular en las redes sociales (Pasquatto et al., 2020).

Sin embargo, la IA es también una espada de doble filo que, por un lado, permite mejorar las tareas humanas en el entorno digital y detectar de forma automatizada textos, imágenes y videos manipulados o contenidos generados por *bots*, pero que en otras manos puede ser una herramienta para la difusión de desinformación a gran escala (Kertysova, 2018). Las advertencias sobre sus potenciales peligros abarcan también nuevas vulnerabilidades, como la preocupación por la veracidad de los contenidos, la ética y la transparencia en el uso de los datos, la intromisión en la vida privada o su contribución al refuerzo de las brechas sociales existentes (Hansen et al., 2017; Brundage et al., 2018; Pihlajarinne & Alén-Savikko, 2022).

Los propios profesionales de la información no son ajenos a este debate. Por un lado, la IA y la automatización de contenidos pueden, al menos potencialmente, aproximar a los periodistas a la esencia de la profesión, liberándolos de las tareas más rutinarias y de menor valor añadido (Wu, Tandoc & Salmon, 2019), lo que a su vez podría reforzar los aspectos más cognitivos de la profesión, como la creatividad, el pensamiento crítico, la curiosidad o el escepticismo (Thurman, Dörr & Kunert, 2017; Guzman & Lewis, 2020). Por su parte, para los medios de comunicación la IA puede ser la llave de acceso a modelos de negocio que mejoren su viabilidad empresarial (Lindén & Tuulonen, 2019) o que les ayuden a reducir costes.

Sin embargo, las nuevas oportunidades no pueden disociarse de los riesgos que implica su implantación. Así, la IA enfrenta a los periodistas y los medios a la irrupción de nuevos agentes generadores de contenido que exceden la función mediadora a la que se había limitado la tecnología (Guzman & Lewis,

2020), lo que puede cuestionar su capital simbólico como mediadores entre la realidad y la ciudadanía, y debilitar su autoridad epistémica (Peña et al., 2023).

En el ámbito laboral, la IA también genera inquietud entre los profesionales de la comunicación por la pérdida de puestos de trabajo que puede suponer su implantación (Kim & Kim, 2017) o por su contribución al incremento de la precariedad profesional (López-Jiménez & Ouariachi, 2020).

En este contexto, esta investigación se plantea analizar el discurso de los periodistas de los medios de comunicación españoles ante la eclosión de las herramientas de inteligencia artificial generativa y su potencial impacto en el riesgo de desinformación. A partir de los resultados de análisis previos sobre la percepción de los periodistas sobre la IA en el periodismo, tanto a nivel español (Calvo-Rubio & Ufarte-Ruiz, 2020; De-Lara-González, García-Avilés, Arias-Robles, 2022; Noain-Sánchez, 2022) como internacional (Jung et al., 2017; Kim & Kim, 2018; Stenborn, Wiggberg & Norlund, 2021; Haan et al., 2022; Soto-Sanfiel et al., 2022), este estudio analiza su influencia en la desinformación a partir de la popularización de aplicaciones de inteligencia artificial generativa como ChatGPT, Bard o Dall-E.

El caso español tiene especial relevancia en el estudio de la interrelación entre IA y desinformación puesto que, a nivel europeo, son los ciudadanos que declaran estar más expuestos a la desinformación y, a su vez, los que muestran menos confianza en su capacidad de detectar las informaciones falsas (Comisión Europea, 2018). Asimismo, los ciudadanos piensan que la responsabilidad para combatir la desinformación recae principalmente sobre los medios de comunicación y los periodistas (Amoedo, Vara-Miguel & Negredo, 2018), lo que refuerza la responsabilidad de los profesionales para combatir los desórdenes informativos.

A partir de este objetivo principal, esta investigación se plantea las siguientes preguntas de investigación.

- RQ1. ¿Cuál ha sido el sentido del discurso de los periodistas españoles sobre la irrupción de la inteligencia artificial generativa?
- RQ2. ¿Qué impacto prevén los periodistas que tendrá la IA generativa?
- RQ3. ¿Cuál es el efecto percibido de la IA sobre la desinformación?

2. Metodología

Para medir la presencia y el sentido del discurso de los periodistas sobre la inteligencia artificial generativa, se ha analizado la presencia del término IA, de sus sinónimos funcionales y sus aplicaciones en las cuentas de Twitter de 420 periodistas españoles de información general.

La elección de Twitter se basa, por un lado, en su popularidad como altavoz a través del que los perso-

najes públicos –en particular en el ámbito de la política– y los periodistas manifiestan sus opiniones. En España, esta red social cuenta con 4,2 millones de cuentas registradas, de las que 796.385 están activas (Social Media Family, 2022). El interés de Twitter también radica en haber sido la plataforma en la que se han basado numerosas herramientas para combatir la desinformación gracias a su API abierta (Míguez-González, Martínez-Rolán & García-Mirón, 2023), aunque los recientes cambios en su política de acceso han cuestionado la viabilidad de estos proyectos.

El periodo analizado ha comprendido seis meses, desde el 1 de noviembre de 2022 hasta el 30 de abril de 2023. Para la definición de la muestra se ha considerado la fecha de lanzamiento de la aplicación ChatGPT, el 30 de noviembre de 2022, como detonante del inicio de la visibilidad y el debate sobre las aplicaciones de inteligencia artificial generativa.

En total, los periodistas y los medios de comunicación seleccionados han publicado durante el periodo analizado 411.847 tuits. La descarga y el análisis inicial de los datos se ha realizado mediante la librería de Python Snsraper, la herramienta de análisis semántico WORDij (Danowski, 2013) y el software 4CAT (Peeters & Hagen, 2022).

De los 420 periodistas analizados, 118 (28,1%) han publicado algún contenido relacionado con la IA durante los seis meses de la muestra. De ellos 87 son hombres y 31 mujeres. Tras realizar el filtrado de los mensajes según los términos de búsqueda seleccionados y depurar los errores, la muestra final la han compuesto 778 mensajes.

Las categorías de análisis se han fijado tras un estudio preliminar o exploratorio sobre un 10% de los mensajes:

- Actitud positiva ante la IA: (1) Tecnología con potencial, (2) Mejora de la productividad, (3) Eliminación de restricciones de uso, (4) Otros.

- Actitud neutra ante la IA: (1) Textos enunciativos, (2) Reto o desafío, (3) Ambivalencia, tiene potencial, pero también peligro, (4) No toda la desinformación proviene de IA, (5) Otros.
- Actitud negativa ante la IA: (1) Superioridad humana (creatividad, etc.), (2) Tecnología insuficientemente desarrollada, (3) Peligro de desinformación, contenidos falsos, (4) Mensajes alarmistas, (5) Necesidad de desarrollar normas, regulación, (6) Paro, precariedad laboral, (7) Impacto en el medio ambiente, (8) Otros.

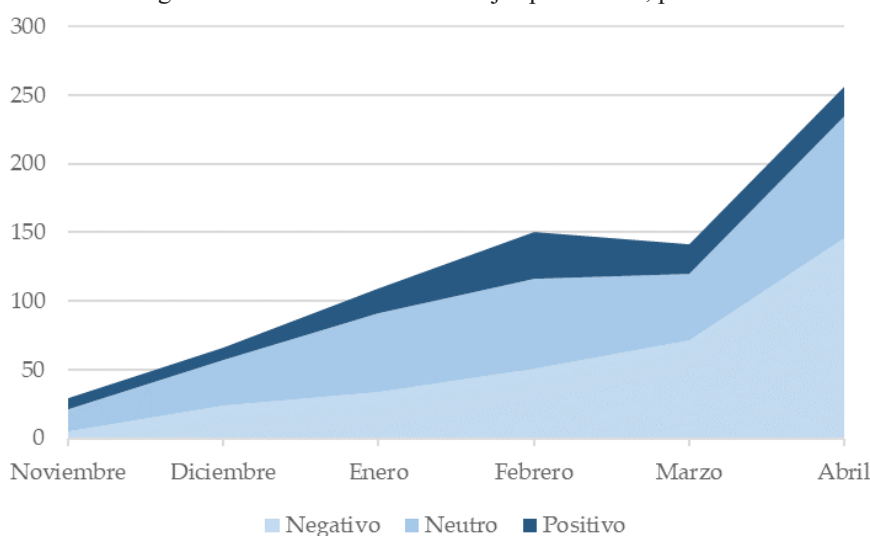
Los mensajes se han etiquetado manualmente para minimizar el factor de riesgo de un análisis automatizado.

3. Resultados

3.1. Difusión

El análisis de los tuits de los periodistas en sus cuentas personales de Twitter indica, en primer lugar, que la presencia de la IA ha sido relativamente limitada en el discurso global y solo algo más de uno de cada cuatro periodistas ha publicado algún contenido referido a esta temática durante los seis meses del estudio. De los 118 periodistas que han mencionado la inteligencia artificial en sus tuits, el 34,7% (n=41) ha publicado un solo mensaje y el 21,7% dos (n=26). En el extremo contrario, 15 de ellos han dedicado al menos diez mensajes a la IA en sus cuentas personales, entre ellos algunos de los profesionales con más de cien mil seguidores en Twitter, como Ana Pastor, Ignacio Escolar, Fernando Berlín, Emilio Doménech o Casimiro García Abadillo. En conjunto, el promedio de tuits publicados ha ascendido a 5,8.

Figura 1. Evolución de los mensajes publicados, por tono



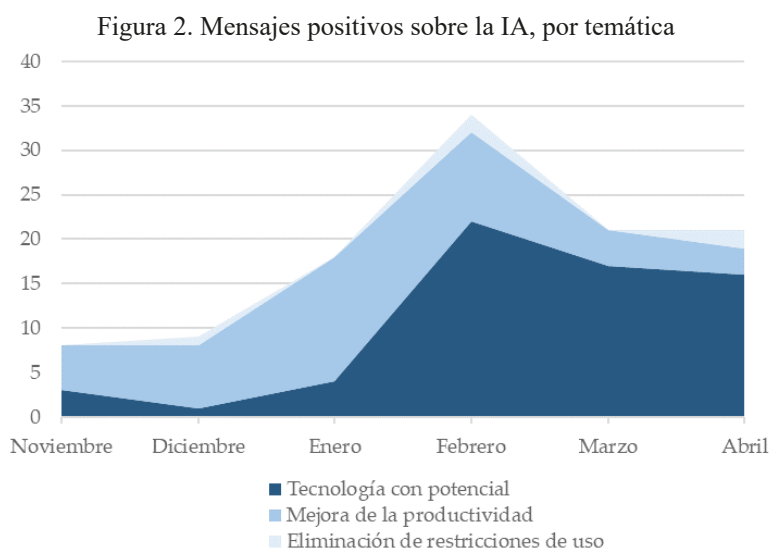
Fuente: Elaboración propia

En la evolución de los mensajes (Figura 1) pueden apreciarse los hitos principales durante el periodo analizado: el lanzamiento de ChatGPT, el 30 de noviembre de 2022 y la publicación el 22 de marzo de la carta en la que mil expertos –entre ellos personalidades como Elon Musk o Steve Wozniak– pedían una moratoria de seis meses en el entrenamiento de los sistemas de inteligencia artificial por tratarse de una “potencial amenaza para la humanidad” (Future of Life Institute, 2023).

En el conjunto del periodo, el discurso general de los periodistas sobre la IA se ha caracterizado por la cautela ante la nueva tecnología. En total, el 44,1% de los mensajes (n=331) ha tenido un tono negativo, frente al 14,8% (n=111) que ha tenido sentido positivo y el 41,1% de tuits neutros (n=309).

3.2. Potencialidades y riesgos

Entre los mensajes favorables a la IA –algo menos de uno de cada siete de los tuits publicados durante el periodo analizado– destacan los que elogian su potencial (56,8%, n=63) (Figura 2). Se subraya en ellos la capacidad que tiene esta tecnología para aportar nuevas funcionalidades, generar contenidos y, en definitiva, su potencial técnico. El segundo rasgo positivo más recurrente que se atribuye a la IA es el incremento de la productividad y las nuevas oportunidades que ofrece a las empresas y los trabajadores para mejorar sus puestos de trabajo (38,7%, n=43). Por último, aunque a mucha distancia de los anteriores, un reducido número de mensajes reivindica la eliminación de restricciones de uso de la IA y el pleno desarrollo de su potencial (4,5%, n=5).



Fuente: Elaboración propia

En la secuencia temporal, los mensajes con tono positivo han perdido fuerza con el paso del tiempo. Si en noviembre de 2022 ascendían 27,6% del total, en el último mes de la muestra –abril de 2023– este porcentaje había descendido hasta el 8,2%. Influye sin duda en ello una primera aproximación a la IA desde la atracción que genera su potencial que redundan en un mensaje algo más tecnocéntrico, pero que da paso poco a poco a una mayor preocupación por los efectos de su implantación y su impacto social.

Los mensajes neutros, por su parte, han experimentado una evolución similar los anteriores (Figura 3). En términos generales, los tuits enunciativos o descriptivos, es decir, aquellos que simplemente mencionan la existencia de una nueva tecnología o la aparición de nuevas aplicaciones sin hacer ningún tipo de valoración han sido los más comunes, un 55,7% (n=171). Las menciones genéricas al reto o desafío que supone la IA, por su parte, han ascendido al 20,8% (n=64) de los mensajes neutros. Por su parte, los tuits ambivalentes, es decir, aquellos que subrayan el potencial de la IA, pero alertan en igual medida sobre sus peligros, han alcanzado el 14,3% (n=44). Finalmente, las afirmaciones que no consideran la inteligencia artificial

como un peligro específico para la desinformación, sino que la ubican dentro de un conjunto global de factores que contribuyen a generar desórdenes informativos, suponen el 9,1% (n=28).

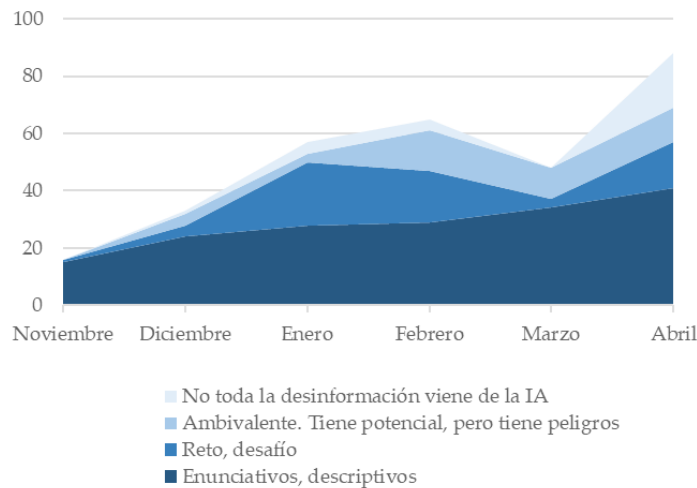
Con el paso del tiempo, los mensajes neutros han dejado de ser simplemente enunciativos o descriptivos para mostrar una mayor interpretación. Si durante los dos primeros meses los contenidos estrictamente informativos supusieron un 79,6% (n=39) de los contenidos considerados neutros, durante los dos últimos meses ese porcentaje ha descendido casi a la mitad, un 42,6% (n=75). Por el contrario, los mensajes ambivalentes, que citan por igual las promesas y las vulnerabilidades de la nueva tecnología, han ascendido paulatinamente durante el mismo periodo. Por ello, sin perder un tono general que se puede calificar como neutral, este tipo de mensajes han pasado de menciones más o menos acríicas en los inicios de las aplicaciones de IA generativa a un tono más analítico a medida que se ha ido conociendo su impacto.

Sin embargo, la carga negativa ha prevalecido en todo momento sobre la positiva en el discurso de los periodistas españoles sobre la IA (Figura 4). Los mensajes que advierten sobre sus peligros triplican a

los que simplemente enumeran sus potencialidades (41,1% frente a 14,8%). Desde las primeras manifestaciones de la IA generativa, los periodistas se han mostrado cautos y han advertido sobre los riesgos que su implantación puede suponer. En este sentido, la preocupación sobre su impacto en el incremento de la desinformación o en la propagación de contenidos falsos ha constituido el tipo de contenido negativo más habitual (36%, n=119), seguido por la reivindicación de la necesidad de una mayor regulación

(17,5%, n=58). También han tenido presencia apelaciones a la superioridad de las cualidades humanas frente a la IA (13,3%, n=44), las advertencias sobre el perjuicio previsto sobre el empleo o el incremento de la precariedad laboral (12,4%, n=41) y la alusión a que se trata de una tecnología insuficientemente desarrollada (10%, n=33). Los mensajes también han dejado ejemplos de contenidos con tono más alarmista (10,6%, n=35) (“ChatGPT: La IA puede socavar la democracia”).

Figura 3. Mensajes neutros sobre la IA, por temática

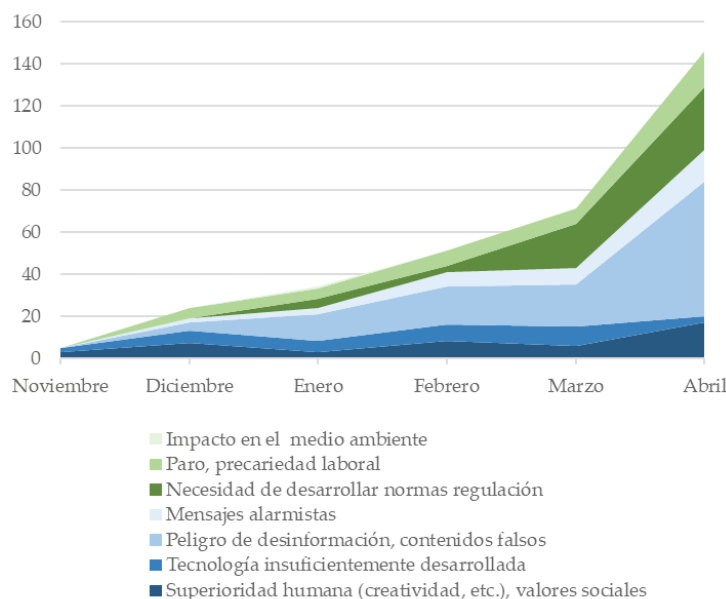


Fuente: Elaboración propia

Si comparamos a nivel global los mensajes positivos y negativos sobre las mismas temáticas, la precaución con la que, de forma general, los periodistas acogen la llegada de la IA generativa queda más patente. Así, frente a los 63 mensajes que mencionan su potencial (8,4% del total), se han publicado 231 que advierten sobre el riesgo de incremento de la desinformación, cuestionan la madurez de la tecnología, reivindican la superioridad humana o se revisten de un tono alarmista (30,8% del total). De la misma for-

ma, frente a los 58 mensajes (7,7%) que solicitan un mayor desarrollo normativo que regule y limite el uso de la IA, solo en 5 tuits (0,7%) se reivindica lo contrario, es decir, poder hacer un uso sin limitaciones de esta tecnología. Algo más parejo es el debate sobre los términos laborales y empresariales, pues frente a los 43 mensajes (5,7%) que afirman que la IA mejorará la productividad, 41 (5,5%) alertan del impacto que tendrá sobre los puestos de trabajo y la precariedad laboral.

Figura 4. Mensajes negativos sobre la IA, por temática



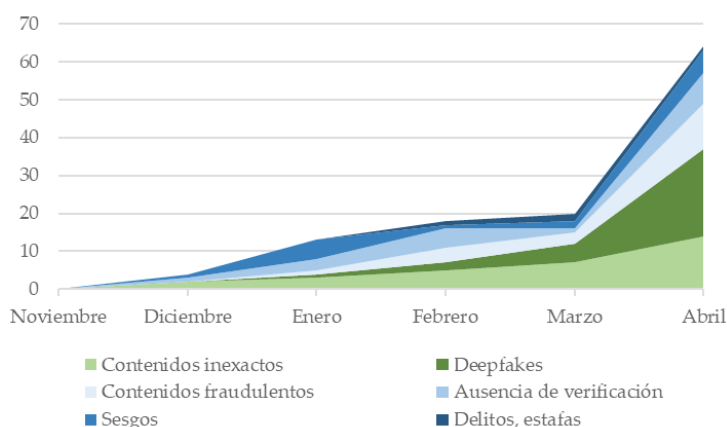
Fuente: Elaboración propia

3.3. El impacto de la IA en la desinformación

Como hemos visto, los contenidos falsos y el riesgo de desinformación constituyen la principal preocupación de los periodistas respecto de la IA generativa (119 mensajes, 15,8%), aunque con un impacto dispar a lo largo del periodo analizado (Figura 5). Si durante los primeros meses esta inquietud apenas obtuvo espacio en su discurso, el impacto de la inteligencia artificial en el riesgo de desinformación comenzó a aflorar de forma mucho más frecuente a partir del mes de abril de 2023, en un conjunto creciente de mensajes de tono negativo.

En esta transformación del discurso tuvieron mucho que ver las pruebas que los propios periodistas comenzaron a hacer con las aplicaciones de IA generativa, que avivaron la preocupación por los contenidos inexactos que pueden obtenerse por esta vía. La información errónea que ofrecen algunos de los resultados de su uso (“La porcelana triturada añade calcio a la leche: el peligro de que la inteligencia artificial genere información sanitaria”) incrementaron la desconfianza sobre la calidad de los resultados y reforzaron la cautela con la que hay que tomarlos. En total, un 26,1% (n=31) de los mensajes relacionados con la desinformación alertaban sobre esta circunstancia, casi la mitad de ellos durante el último mes de la muestra.

Figura 5. Mensajes sobre desinformación



Fuente: Elaboración propia

En segundo lugar, los periodistas manifiestan su inquietud por el modo en el que la IA puede ser una herramienta para crear contenidos de forma fraudulenta (16,8%, n=20). En este apartado, durante el periodo analizado destaca en especial la destitución de la directora de la revista alemana *Die Aktuelle* por publicar en primera página una supuesta entrevista al excampeón del mundo Michael Schumacher, generada por completo con IA y ambiguamente atribuida. En general, la facilidad con la que pueden generarse contenidos que se atribuyan falsamente preocupa a los periodistas (“La inteligencia artificial puede convertirse en una herramienta de plagio masivo... interesante reflexión”).

A una escala de mayor intensidad en los desórdenes informativos, los *deepfakes* o la capacidad de la IA generativa para imitar la realidad o a las personas de forma creíble inquieta también los periodistas (“El gran desafío de la inteligencia artificial: Pronto será imposible distinguir la verdad de la mentira”). En conjunto, estos mensajes alcanzaron el 26,1% de los tuits relacionados con la desinformación. Entre los ejemplos más destacados pueden citarse los casos de las fotografías sobre la supuesta detención de Donald Trump en Nueva York en marzo de 2023, creadas con vocación ilustrativa pero no siempre bien contextualizadas, o una supuesta nueva canción de los cantantes Drake y The Weeknd generada parcialmente mediante IA.

Asimismo, desde las cuentas analizadas los profesionales también han alertado sobre los usos malicio-

sos y delictivos de la IA, y el modo en el que pueden incrementar las vulnerabilidades y la seguridad de los ciudadanos. Su potencial uso con fines criminales se suma a otros contenidos de tono alarmista, publicados en especial durante los meses de marzo y abril de 2023 (“El timo del ‘skimming’, estafas a través del 118 e inteligencia artificial para suplantar la voz de nuestros conocidos”). Pese a ello, estos peligros no directamente relacionados con la actividad periodística tuvieron un impacto proporcionalmente menor (3,4%).

Por último, los periodistas también se han hecho eco de los sesgos mostrados por los contenidos generados por la IA, que pueden contribuir al incremento de brechas sociales ya existentes (de género, raza, clase social, etc.) por utilizar para su entrenamiento de sus modelos bancos de datos parciales o anticuados (“En este programa de inteligencia artificial, cuando das la orden de crear imágenes de gente, por defecto son blancas. La IA también es racista” / “La inteligencia artificial también es machista: no discrimina, pero refuerza los estereotipos de género”). En total 15 mensajes (12,6%) resaltaron este extremo.

4. Conclusiones y discusión

Durante los últimos años, los medios de comunicación han contribuido a difundir la percepción de que la IA pronto estará presente en todos los ámbitos de

nuestra vida (Brennen, Howard & Nielsen, 2018), y la irrupción de la IA generativa ha ofrecido a la ciudadanía herramientas tangibles para percibir dicho cambio.

Ante la eclosión de estas herramientas tecnológicas, el discurso de los periodistas españoles ha evolucionado desde un tono principalmente positivo o enunciativo, centrado describir las potencialidades de la tecnología, a un tono más crítico y analítico, que ha resaltado las anomalías y las disfunciones en los resultados de su uso y las alertas difundidas desde otros ámbitos. En conjunto, los mensajes de tono negativo o de advertencia sobre los riesgos de la implantación de la inteligencia artificial han triplicado a los mensajes positivos.

En general, los periodistas españoles expresan en sus cuentas en Twitter una actitud mayoritariamente cauta sobre la implantación de la IA en la que huyen de visiones tecnocéntricas, y prevalece entre ellos la preocupación por el modo en el que esta tecnología puede afectar a la ciudadanía (De-Lara, García-Avilés & Arias, 2022) y a la propia profesión (Noain-Sánchez, 2022).

Así, los periodistas muestran una mayor inquietud por el impacto social de la IA que por su potencial tecnológico, abogan por una mayor regulación en este ámbito y, a pesar de reconocer su carácter transformador sobre la economía y la empresa, expresan sus recelos sobre el efecto que tendrá sobre el empleo y las condiciones de trabajo en los medios de comunicación. Por ello, la eventual erosión de su rol social

y profesional (Kim & Kim, 2018; López-Jiménez, & Ouariachi, 2020) obliga a repensar el modo en el que se configura la profesión periodística ante la irrupción de esta tecnología (Lewis, Guzman & Schmidt, 2019; Deuze & Beckett, 2022).

Además, para los periodistas el incremento de los desórdenes informativos y la desinformación se erige en el principal peligro ocasionado por el desarrollo de la IA generativa, y se enumeran riesgos como los sesgos en los resultados obtenidos y los contenidos inexactos (Pihlajarinne & Alén-Savikko, 2022), la facilidad para cometer fraude o elaborar *deepfakes* (Westerlund, 2019), o su potencial uso con fines maliciosos y delictivos (Brundage et al., 2018).

Ante la creciente implantación de la inteligencia artificial en el trabajo de los medios de comunicación, el debate sobre su naturaleza dual (Brennen, Howard & Nielsen, 2018), que oscila entre los beneficios que ofrece y las nuevas vulnerabilidades que puede crear su uso, el discurso de los periodistas se ha inclinado en favor de las segundas con el paso del tiempo.

Ante el efecto mitigador, pero también amplificador, que la IA generativa puede tener sobre la desinformación, los profesionales de los medios de comunicación reivindican mayoritariamente aumentar la regulación de la IA (Smuha, 2021), en la línea de lo que ya han comenzado a promover diferentes instituciones y organismos, y defienden los principios editoriales y el control profesional sobre los contenidos (Wu, Tandoc & Salmon, 2019) como las principales señas de identidad del periodismo.

Referencias

- Altunbey Ozbay, F., & Alatas, B. (2020). Fake news detection within online social media using supervised artificial intelligence algorithms. *Physica A: Statistical Mechanics and its Applications*, 540, 123174. <https://doi.org/10.1016/j.physa.2019.123174>.
- Amoedo, A., Vara-Miguel, A., & Negrodo, S. (2018). *Digital News Report.ES 2018*. Pamplona: Servicio de Publicaciones de la Universidad de Navarra. <https://doi.org/10.15581/019.008>
- Beckett, C. (2019). New powers, new responsibilities: A global survey of journalism and artificial intelligence. The London School of Economics and Political Science. <https://tinyurl.com/46nrjsss>
- Brennen, J. S., Howard, P. N., & Nielsen, R. K. (2018). An industry-led debate: how UK media cover artificial intelligence. Reuters Institute for the Study of Journalism. <https://tinyurl.com/mr2va4dr>
- Brennen, J. S., Howard, P. N., & Nielsen, R. K. (2022). What to expect when you're expecting robots: Futures, expectations, and pseudo-artificial general intelligence in UK news. *Journalism*, 23(1), 22-38. <https://doi.org/10.1177/1464884920947535>
- Brundage, M. et al. (2018). The Malicious Use of Artificial Intelligence: Forecasting, Prevention, and Mitigation. <https://tinyurl.com/3ayc5tnw>
- Calvo-Rubio, L. M., & Ufarte-Ruiz, M. J. (2020). Percepción de docentes universitarios, estudiantes, responsables de innovación y periodistas sobre el uso de inteligencia artificial en periodismo. *Profesional de la Información*, 29(1). <https://doi.org/10.3145/epi.2020.ene.09>
- Cea, N., & Palomo, B. (2021). Disinformation Matters. En: Lopez-Garcia, G., Palau-Sampio, D., Palomo, B., Campos-Dominguez, E. & Masip, P. (eds) *Politics of Disinformation* (pp. 7-23). Wiley.
- Clayton, K., Blair, S., Busam, J.A. et al. (2020). Real Solutions for Fake News? Measuring the Effectiveness of General Warnings and Fact-Check Tags in Reducing Belief in False Stories on Social Media. *Political Behaviour* 42, 1073–1095. <https://doi.org/10.1007/s11109-019-09533-0>
- Comisión Europea (2018). *Flash Eurobarometer 464: Fake News and Disinformation Online*. Bruselas: Directorate-General for Communication. <https://tinyurl.com/43p8pnut>
- Comisión Europea (2020). *Libro blanco sobre la inteligencia artificial. Un enfoque europeo orientado a la excelencia y la confianza*. <https://tinyurl.com/mr42435s>
- Danowski, J. A. (2013). *WORDij version 3.0: Semantic network analysis software*. University of Illinois at Chicago.

- De-Lara-González, A., García-Avilés, J.A., & Arias-Robles, F. (2022). Implantación de la Inteligencia Artificial en los medios españoles: análisis de las percepciones de los profesionales. *Textual & Visual Media*, 1(15), 1-16. <https://doi.org/10.56418/txt.15.2022.001>
- Deuze, M.; Beckett, C. (2022). Imagination, Algorithms and News: Developing AI Literacy for Journalism. *Digital Journalism*. <https://doi.org/10.1080/21670811.2022.2119152>
- Flores Vivar, J. M. (2019). Inteligencia artificial y periodismo: diluyendo el impacto de la desinformación y las noticias falsas a través de los bots. *Doxa Comunicación*, 29, 197-212. <https://doi.org/10.31921/doxacom.n29a10>
- Future of Life Institute (22 de marzo de 2023). Pause Giant AI Experiments: An Open Letter. <https://tinyurl.com/y7dt9d5c>
- Guzman A. L., & Lewis S. C. (2020). Artificial intelligence and communication: A human-machine communication research agenda. *New Media & Society* 22(1). <https://doi.org/10.1177/1461444819858691>
- Haan, Y., van den Berg, E., Goutier, N., Kruikemeier, S., & Lecheler, S. (2022). Invisible Friend or Foe? How Journalists Use and Perceive Algorithmic-Driven Tools in Their Research Process. *Digital Journalism*. <https://doi.org/10.1080/21670811.2022.2027798>
- Hansen, M., Roca-Sales, M., Keegan, J., & King, G. (2017). Artificial Intelligence: Practice and Implications for Journalism. Columbia Journalism School. <https://doi.org/10.7916/D8X92PRD>
- Holton, A. E., & Belair-Gagnon, V. Strangers to the Game? Interlopers, Intralopers, and Shifting News Production. *Media and Communication*, 6(4), 70-78. <https://doi.org/10.17645/mac.v6i4.1490>
- Jung, J., Song, H., Kim, Y., Im, H., & Oh, S. (2017). Intrusion of software robots into journalism: The public's and journalists' perceptions of news written by algorithms and human journalists. *Computers in Human Behavior*, 71, 291-298, <https://doi.org/10.1016/j.chb.2017.02.022>
- Kertysova, K. (2018). Artificial Intelligence and Disinformation. *Security and Human Rights*, 29(1-4), 55-81. <https://doi.org/10.1163/18750230-02901005>
- Kim, D., & Kim, S. (2017). Newspaper companies' determinants in adopting robot Journalism. *Technological forecasting and social change*, 117, 184-195. <https://doi.org/10.1016/j.techfore.2016.12.002>
- Kim, D., & Kim, S. (2018). Newspaper journalists' attitudes towards robot Journalism. *Telematics and Informatics*, 35(2), 340-357. <https://doi.org/10.1016/j.tele.2017.12.009>
- Lewis, S. C., Guzman A. L., & Schmidt, T. R. (2019). Automation, Journalism, and Human-Machine Communication: Rethinking Roles and Relationships of Humans and Machines in News. *Digital Journalism*, 7(4), 409-427. <https://doi.org/10.1080/21670811.2019.1577147>
- Lindén, C. G., & Tuulonen, H. (2019). *News automation: The rewards, risks and realities of 'machine journalism'*. WAN-IFRA Report. <https://tinyurl.com/yc24pmxc>
- López-Jiménez, E., & Ouariachi, T. (2020). An exploration of the impact of artificial intelligence (AI) and automation for communication professionals. *Journal of Information, Communication and Ethics in Society*, 19(2), 249-267. <https://doi.org/10.1108/JICES-03-2020-0034>
- Marwick, A., & Lewis, R. (2017). *Media Manipulation and Disinformation Online*. Data & Society Research Institute. <https://tinyurl.com/2a9samnf>
- Masip, P., Ruiz-Caballero, C., & Suau, J. (2019). Active audiences and social discussion on the digital public sphere. Review article. *El profesional de la información*, 28(2), e280204. <https://doi.org/10.3145/epi.2019.mar.04>
- Noain-Sánchez, A. (2022). Addressing the Impact of Artificial Intelligence on Journalism: The perception of experts, journalists and academics. *Communication & Society*, 35(3), 3. <https://doi.org/10.15581/003.35.3.105-121>
- Pasquetto, I. et al. (2020). Tackling misinformation: What researchers could do with social media data. *Harvard Kennedy School (HKS) Misinformation Review*, 1(8), 1-15. <https://tinyurl.com/233v5h4w>
- Peeters, S., & Hagen, S. (2022). The 4CAT Capture and Analysis Toolkit: A Modular Tool for Transparent and Traceable Social Media Research. *Computational Communication Research*, 4(2). <https://tinyurl.com/3yzz22y3>
- Peña-Fernández, S., Meso-Ayerdi, K., Larrondo-Ureta, A., & Díaz-Noci, J. (2023). "Without journalists, there is no journalism: the social dimension of generative artificial intelligence in the media". *Profesional de la información*, 32(2), e320227. <https://doi.org/10.3145/epi.2023.mar.27>
- Pihlajarinne, T. & Alén-Savikko, A. (2022). *Artificial Intelligence and the Media. Reconsidering Rights and Responsibilities*. Edward Elgar Publishing. <https://tinyurl.com/2zdbrepu>
- Posetti, J., & Matthews, A. (2018). *A short guide to the history of 'fake news' and disinformation*. International Center for Journalists, 7, 1-19. <https://tinyurl.com/bddk27te>
- Smuha, N. A. (2021). From a 'race to AI' to a 'race to AI regulation': regulatory competition for artificial intelligence. *Law, Innovation and Technology*, 13(1), 57-84. <https://doi.org/10.1080/17579961.2021.1898300>
- Social Media Family (2022). *VIII Informe sobre el uso de las redes sociales en España*. <https://tinyurl.com/5fsd2crf>
- Soto-Sanfiel, M. T., Ibíti A., Machado, M., Marín-Ochoa, B. E., Mendoza-Michilot, M., Rosell-Arce, C. G., & Angulo-Brunet, A. (2022). In Search of the Global South: Assessing Attitudes of Latin American Journalists to Artificial Intelligence in Journalism. *Journalism Studies*, 23(10), 1197-1224. <https://doi.org/10.1080/1461670X.2022.2075786>
- Stenborn, A., Wiggberg, M., & Norlund, T. (2021). Exploring Communicative AI: Reflections from a Swedish Newsroom. *Digital Journalism*. <https://doi.org/10.1080/21670811.2021.2007781>
- Sunstein, C. R. (2017). *#Republic: Divided Democracy in the Age of Social Media*. Princeton University Press.

- Waisbord, S. (2018) Truth is What Happens to News. *Journalism Studies*, 19:13, 1866-1878. <https://doi.org/10.1080/1461670X.2018.1492881>
- Wardle, C., & Derakhshan, H. (2017). *Information disorder: Toward an interdisciplinary framework for research and policy making*. Consejo de Europa.
- Westerlund, M. (2019). The Emergence of Deepfake Technology: A Review. *Technology Innovation Management Review*, 9(11), 40-53. <http://doi.org/10.22215/timreview/1282>
- Wu, S., Tandoc, E. C., & Salmon, C. T. (2019). Journalism Reconfigured. *Journalism Studies*, 20(10), 1440-1457. <https://doi.org/10.1080/1461670X.2018.1521299>
- Zhuk, D., Tretiakov, A., Gordeichuk, A., & Puchkovskaia, A. (2018). Methods to Identify Fake News in Social Media Using Artificial Intelligence Technologies. In D. Alexandrov, A. Boukhanovsky, A. Chugunov, Y. Kabanov, & O. Koltsova (Eds) *Digital Transformation and Global Society. DTGS 2018. Communications in Computer and Information Science*, 858. Springer, Cham. https://doi.org/10.1007/978-3-030-02843-5_36

Simón Peña-Fernández. Profesor pleno en el departamento de Periodismo de la UPV/EHU. Licenciado en Periodismo y en Comunicación Audiovisual por la Universidad de Navarra y Doctor en Ciencias de la Información por la UPV/EHU, ha publicado 70 artículos en revistas académicas. Todas estas publicaciones han ido siempre asociadas a la participación continuada en una treintena de proyectos de investigación financiados en convocatorias públicas competitivas, entre ellos, ocho proyectos europeos (Horizon 2020 y Erasmus+) y cuatro proyectos del Plan Nacional. También ha liderado ocho contratos de investigación y transferencia relevantes con entidades públicas y ha participado en cinco proyectos de innovación educativa. ORCID: <https://orcid.org/0000-0003-2080-3241>

Urko Peña-Alonso. Investigador contratado en la UPV/EHU, forma parte del proyecto *DISARGUE: Argumentation-based Explainable Detection and Mitigation of Misbehaviour in Social Media*. Es graduado en Comunicación y obtuvo el Máster en Comunicación Social. Sus principales líneas de investigación son las redes sociales, la desinformación y la moderación de contenidos. ORCID: <https://orcid.org/0000-0002-9214-5906>

Maidier Eizmendi Iraola. Profesora en el departamento de Periodismo de la Universidad del País Vasco (UPV/EHU). Licenciada en Periodismo y Antropología Social y Cultural, es doctora en Comunicación Social y obtuvo el Máster en Igualdad de Mujeres y Hombres. Cuenta con una trayectoria de 16 años como periodista en el periódico Gara. Como profesora de la UPV/EHU, ha impartido docencia de Fotoperiodismo, Redacción Informativa en prensa, Edición en prensa, Redacción y creación de contenidos web y Edición y Producción multimedia. Miembro del equipo de trabajo del grupo de investigación Gureiker, ha publicado varios artículos en revistas académicas y ha presentado comunicaciones en congresos nacionales e internacionales. ORCID: <https://orcid.org/0000-0002-5894-8238>