

Implicaciones éticas del uso de IA en verificación. Percepciones de verificadores y académicos en España

Amaya Noain-Sánchez

Universidad Rey Juan Carlos ✉ 

Leticia Rodríguez-Fernández

Universidad de Cádiz ✉ 

Marina Santín-Durán

Universidad Rey Juan Carlos ✉ 

<https://dx.doi.org/10.5209/esmp.106640>

Recibido: 19 de diciembre de 2025 • Aceptado: 2 de marzo de 2026

ES Resumen. Se explora cómo perciben los profesionales de la verificación, expertos y académicos el uso de la inteligencia artificial (IA) en los procesos de comprobación informativa, centrándonos en su utilidad, implicaciones éticas y la existencia de protocolos que regulen su uso. Se adopta un enfoque cualitativo mediante 16 entrevistas semiestructuradas a verificadores y académicos españoles, realizadas entre enero y junio de 2025. El análisis con Atlas.ti, aplica una codificación inductiva para identificar patrones y percepciones sobre el uso de IA en verificación. Se observa que la IA mejora la eficiencia en la fase previa a la verificación y posterior distribución de contenidos verificados, pero en la fase de comprobación su autonomía es limitada y los sistemas automatizados presentan baja fiabilidad. La falta de neutralidad de las herramientas comerciales usadas en verificación provoca reticencias en algunos de los entrevistados que demandan un desarrollo más transparente de la tecnología. Además, se identifican riesgos asociados a sesgos algorítmicos, falta de conjuntos de datos adecuados para entrenar los programas, dificultad para asegurar la fiabilidad de las fuentes y opacidad en los procesos, lo que amplía las preocupaciones éticas en torno a la veracidad de los resultados. El estudio concluye que la IA debe entenderse como recurso complementario bajo el modelo *human-in-the-loop*, sin sustituir el juicio profesional. En relación con la demanda de transparencia, destaca la necesidad de impulsar el principio de explicabilidad y desarrollar códigos éticos específicos que garanticen trazabilidad, neutralidad y protección frente a sesgos, en consonancia con estándares internacionales y demandas sociales de confianza en la información.

Palabras clave: Desinformación, ética, verificación de hechos, inteligencia artificial, deontología.

ENG Ethical Implications of Artificial Intelligence in Fact-Checking: Insights from Fact-checkers and Scholars in Spain

Abstract. This study examines how fact-checking professionals, experts, and scholars perceive the use of artificial intelligence (AI) in information verification processes, focusing on its utility, ethical implications, and the existence of codes governing its application. A qualitative approach was adopted through 16 semi-structured interviews with Spanish fact-checkers and academics conducted between January and June 2025. The analysis carried out using Atlas.ti, employed inductive coding to identify patterns and perceptions regarding AI in verification. Findings indicate that AI enhances efficiency in the pre-verification stage and in the subsequent dissemination of verified content. However, during the core verification phase, its autonomy remains limited, and automated systems exhibit low reliability. The lack of neutrality in commercial tools used for verification generates skepticism among interviewees, who call for greater transparency in technological development. Additional risks include algorithmic bias, insufficient datasets, challenges in ensuring source reliability, and opacity in processes, all of which heighten ethical concerns about the accuracy of results. The study concludes that AI should be understood as a complementary tool within a human-in-the-loop model, rather than a substitute for professional judgment. Furthermore, in response to demands for transparency, the research underscores the need to promote the principle of explainability and to develop specific ethical codes ensuring traceability, neutrality, and protection against bias, aligned with international standards and societal expectations for trustworthy information.

Keywords: Disinformation, ethics, fact-checking, artificial intelligence, deontology.

Cómo citar: Noain-Sánchez, A., Rodríguez-Fernández, L. y Santín-Durán, M. (2026). Implicaciones éticas del uso de IA en verificación. Percepciones de verificadores y académicos en España. *Estudios sobre el Mensaje Periodístico*, 32(3), 515-527. <https://dx.doi.org/10.5209/esmp.106640>

1. Introducción

Las herramientas de inteligencia artificial (IA), tecnologías informáticas que simulan la inteligencia humana (Sanguinetti, 2023) permiten automatizar tareas rutinarias en el periodismo y facilitan la investigación y/o análisis de grandes volúmenes de datos, por lo que su inclusión en la verificación de información se erige como una funcionalidad con enorme proyección (Beckett y Yaseen, 2023). Las plataformas y áreas de verificación de datos o *fact-checking*, organizaciones integradas por profesionales encargados de contrastar la información, bien desde proyectos anclados a grupos de comunicación o como iniciativas independientes ligadas a entornos alternativos o colaborativos (Graves, 2018) han incorporado en los últimos años la IA en sus procesos. Aun cuando muchas de estas organizaciones operan con estructuras organizativas y recursos escasos (Tejedor y Sancho, 2023; Vizoso y Vázquez, 2019) incluir IA en los procesos de verificación les permite hacer frente al notable incremento, tanto en magnitud, como en celeridad, con que la desinformación circula en los entornos digitales (Vosoughi et al., 2018). La IA se aplica a la monitorización y detección de afirmaciones susceptibles de verificación (*detection*), recopilación de información necesaria para contrastar datos (*reporting*) y evaluación de posibles falsedades (*debunking*) (Guo et al., 2022). Su implementación posee además una vertiente cualitativa: parte de esta desinformación, un 10 %, ya está creada con IA (EDMO, 2025) y para análisis se necesitan herramientas capaces de observar patrones y trazas que escaparían al ojo humano.

Empero, su uso no está exento de conflictos éticos. Para O'Neil (2016) o Noble (2018) la IA no es neutral: refleja y amplifica sesgos presentes en los datos y en los sistemas que la entrenan, por lo que la automatización de procesos informativos puede reproducir desigualdades y favorecer dinámicas de polarización. El hecho de que estas tecnologías operen con lógicas opacas obliga a reflexionar sobre su aplicación y si integra los principios de la práctica periodística. La ética de la información «se ha convertido en un imperativo en la era algorítmica» (Floridi, 2019).

La percepción del usuario también revela esta doble vertiente. Según el Monitor de IA de IPSOS, el 69 % usa la IA en sus búsquedas y el 54 % confía más en su imparcialidad que en la conducta humana, sin embargo, la mayoría prefiere el contenido realizado por humanos y el 79 % está de acuerdo (40 % muy de acuerdo) en que las empresas que la usen lo declaren (IPSOS, 2025).

Consciente de ello, la UE aprobó el Reglamento (UE) 2024/1689 de Inteligencia Artificial (RIA) (BOE, 2024) que regula el desarrollo e implementación de sistemas de IA en diversos sectores, incluido el mediático, promoviendo la autorregulación ética dentro de los propios medios, mediante códigos de conducta y principios editoriales adaptados al entorno digital. El hecho de que se trate de un Reglamento, con la obligatoriedad que este conlleva, ya refleja que la problemática no es un tema menor. Tras este marco general, no obstante, existe una ausencia de protocolos que establezcan marcos de uso responsables y seguros, algo que puede lastrar el trabajo de los verificadores, en tanto que la exigencia de rigor,

imparcialidad, rendición de cuentas y transparencia en la metodología constituye el pilar que fundamenta su ejercicio profesional (Singer, 2021).

Este estudio, planteado a modo prospectivo, pretende recabar la visión de los verificadores y expertos españoles, identificando las herramientas de IA empleadas en la verificación de contenidos y analizando su eficacia/necesidad, nivel de integración, autonomía y necesidad de supervisión en dichos procesos. Su principal aportación, es que se pone el foco en las implicaciones éticas que plantea su uso y evalúa las necesidades de establecer guías profesionales o códigos éticos que enmarquen su uso en verificación.

2. Marco teórico

2.1. La introducción de la IA en verificación

La integración de la IA en las plataformas de *fact-checking* presenta dos enfoques principales: a) el *fact-checking* completamente automatizado o computacional, que permite la comprobación automática de la veracidad de una información «para clasificarla en alguna de las categorías comúnmente usadas por los *fact-checkers* humanos: verdad, media verdad, engañosa, falsa, etc.» (García-Marín, 2022, p. 8); y b) el enfoque híbrido, *human-in-the-loop* (Milosavljević y Vobič, 2019) o «*fact-checking* asistido» (La-Barbera et al., 2022) que permite optimizar y simplificar acciones complementarias al trabajo de comprobación humana.

La procedencia de estas herramientas es heterogénea. Aunque son muchos los desarrollos comerciales que pueden usar las organizaciones también existen herramientas gratuitas: motores de búsqueda especializados en la recuperación de contenidos verificados (como FactCheck Explorer, de Google); motores de búsqueda inversa de imágenes (TinEye); para análisis forense de imágenes (Forensically); verificación audiovisual (inVID WeVerify) o de *deepfakes* (Deepware) entre otras (García-Marín, 2025). Asimismo, los *fact-checkers* cuentan con herramientas propias que desarrollan para responder a necesidades específicas de verificación y distribución. Es el caso de Full Fact, que introdujo en 2016 el aprendizaje automático para mejorar y ampliar la verificación de datos y que ha desarrollado Full Fact AI, un sistema de herramientas usado por 40 organizaciones de verificación que trabajan en tres idiomas en 30 países (Full Fact, 2025).

En España, podemos citar la herramienta modular desarrollada dentro del proyecto IVERES de RTVE que, igualmente, pone a disposición de otros usuarios dos cajas de herramientas gratuitas y *chatbots* como los de *Maldita.es* y *Newtral* (Moreno-Gil, 2024) siendo estos últimos, los *chatbots*, las herramientas más empleadas las entidades de verificación (Canavilhas e Ito, 2025). La capacidad de estos *chatbots* para reducir la creencia en conspiraciones generando argumentos persuasivos, fundamentados en pruebas y contraargumentaciones personalizadas en formato conversacional se ha demostrado en proyectos como Debunkbot creado por el MIT (Costello et al., 2024) e incluye la capacidad de crear explicaciones más allá de los bulos individuales (Gelado-Marcos et al., 2025; Ventura-Salom y Gelado-Marcos, 2025).

Aunque no todos los verificadores tienen recursos suficientes para desarrollar sus propias aplicaciones, el hecho de que trabajen agrupados en consorcios internacionales permite a entidades más pequeñas acceder a herramientas tecnológicas de frontera (García-Marín, 2025; Sánchez-González *et al.*, 2022). Asimismo, el esfuerzo por dotarse de recursos tecnológicos propios se extiende también hacia iniciativas de mayor alcance y aplicabilidad que terminan convirtiéndose en recursos compartidos. Es el caso de ClaimHunter, que detecta afirmaciones realizadas por políticos o Claim-Check, que complementa a la anterior y permite detectar cuándo una afirmación ya verificada reaparece de nuevo en el debate público (Larraz *et al.*, 2023). Ambas fueron desarrolladas por Newtral, pero son empleadas por otros verificadores como Verificat o Maldita (Alonso-González y Sánchez Gonzales, 2024).

2.2. Retos éticos y deontológicos

Son varios los organismos internacionales, centros de investigación, medios de comunicación y organizaciones profesionales del sector, liderados principalmente por medios estadounidenses y, sobre todo, europeos, que han impulsado iniciativas para regular un uso responsable de la IA. En Estados Unidos, el *Código de principios de la International Fact-Checking Network* (IFCN, 2024), perteneciente al Instituto Poynter y en Europa, el *Código de buenas prácticas de la European Fact-Checking Standards Network* (EFCSN, 2022) constituyen las dos referencias de las que son signatarias un gran número de verificadores. El código de la IFCN, por ejemplo, recoge aspectos ligados a la imparcialidad, el no partidismo y el compromiso con la transparencia, tanto en metodología y uso de fuentes como en la financiación. El de la EFCSN, por su parte, se desarrolla sobre tres pilares: metodología, ética y transparencia, abordando aspectos que van desde la calidad de las fuentes usadas, la imparcialidad o el apartidismo, hasta la transparencia financiera u organizacional. Investigaciones centradas en el cumplimiento de estos códigos (Moreno-Gil y Salgado-de Dios, 2023, Torres-Toukoumidis *et al.*, 2021) identifican un alto grado de compromiso con la transparencia metodológica, pero no así, cuando se trata de rendir cuentas sobre la financiación.

En Europa, la redacción del *Código de Buenas Prácticas para la IA de uso general* (Comisión Europea, 2025) dirigido a desarrolladores, fue objeto de controversia y Reporteros Sin Fronteras (RSF) se retiró del proceso de elaboración al considerar que no ofrecía suficientes garantías para la protección del periodismo. Como reacción, RSF impulsó junto a 16 organizaciones aliadas la *Carta de París sobre IA y Periodismo* (2023), que establece principios éticos específicos para todos los periodistas, redacciones y medios de comunicación. Su desvinculación de la citada normativa europea pone de relieve la excesiva influencia de la industria tecnológica en la definición de estándares éticos, especialmente en lo relativo al impacto de la IA sobre la libertad de información y el ejercicio del periodismo. El hecho ilustra que la elaboración de códigos éticos periodísticos en el ámbito de la IA presenta una notable complejidad por la diversidad de agentes involucrados entre los que persisten tensiones significativas.

En España destaca la pionera guía del Consell de la Informació de Catalunya (CIC) con ocho recomendaciones orientadas a preservar los valores fundamentales del periodismo en entornos automatizados. Entre otras, destacan la necesidad de garantizar la calidad y representatividad de los datos utilizados por los algoritmos, la transparencia en su funcionamiento, la protección de la privacidad de los usuarios y el pluralismo informativo (Ventura, 2021).

Entre los principios, la transparencia, aunque pueda no ser siempre entendida de manera idéntica, es un requisito sobre el que parece haber cierta unanimidad en el debate académico y profesional (Sanahuja-Sanahuja y López-Rabadán, 2025). Dado que «resulta fundamental que estas entidades logren la confianza y credibilidad por parte de las audiencias» (García-Marín, 2024:3), los verificadores no deben asumir «una posición neutral y no partidista en las diferentes fases del proceso de verificación» (García-Marín *et al.*, 2023, p. 211) a la vez que deben garantizar unos estándares de transparencia sobre su propiedad y financiación. Cuartielles *et al.* (2024) evidencian que entidades especializadas en verificación como *Maldita*, *Newtral*, *EFE Verifica*, *Verifica TVE* o *Verificat* consideran este principio como fundamental, especialmente, en lo que respecta a sus metodologías, mientras que Sanahuja-Sanahuja y López-Rabadán (2022) identifican que *Maldita*, *Newtral*, *EFE Verifica* y *Verifica RTVE* hacen referencia a cómo la IA influye en sus procesos informativos. Destaca el caso de *Verifica RTVE* que publica en su web todas las herramientas utilizadas con la intención de apostar por «la alfabetización mediática y digital y por el empleo de herramientas de investigación de libre acceso, para que cualquier ciudadano pueda investigar por su cuenta» y verificar imágenes, videos y datos (Huwyler, 2022).

Empero, es especialmente relevante la prácticamente ausencia, entre los verificadores españoles, de guías específicas sobre la aplicación de la IA en sus procesos de verificación. Tan sólo *Newtral* ha publicado un texto breve a modo de guía ética para el uso de la IA generativa, donde menciona una aplicación responsable destinada a «ayudar a los verificadores en sus tareas de investigación, comprobación y contraste de datos, en los procesos de escucha y monitorización, con el objetivo de que cada redactor pueda ser más eficiente. Pero no para sustituirlos por máquinas» (*Newtral*, 2024) y siempre, respetando los derechos de autor. Se comprometen a no utilizar algoritmos opacos de terceros sino a producirlos ellos mismos y supervisar su desarrollo para controlar su arquitectura y contenidos. Así como a publicar su metodología de verificación, explicando procedimientos, tecnologías y funcionamiento de los algoritmos que entrenan y utilizan (*Newtral*, 2024).

3. Objetivos y metodología

El objetivo central de esta investigación, conocer las percepciones de profesionales de la verificación, expertos y académicos sobre el uso de herramientas de IA en tareas propias de comprobación de la información, dedicando especial atención a las implicaciones éticas que genera el uso de estas herramientas, se desgrana en cuatro objetivos concretos: O1. Averiguar qué herramientas de IA se están aplicando

en tareas propias de la comprobación de contenidos. O2. Describir el nivel de integración, autonomía y supervisión que caracteriza a estas herramientas aplicadas al proceso de verificación. O3. Conocer qué implicaciones éticas vinculadas al uso de las herramientas de IA en tareas de comprobación observan tanto profesionales de la verificación como investigadores. O4. Averiguar cómo perciben los entrevistados la necesidad de incorporar guías profesionales o códigos éticos para una correcta utilización de las herramientas de IA en tareas de verificación.

Para abordar estos objetivos se optó por un enfoque cualitativo con entrevistas semiestructuradas. El carácter exploratorio de esta investigación se justifica en la novedad que aún supone la integración de la IA en tareas propias del periodismo. García-Avilés *et al.*, 2018 señalan la adecuación de esta aproximación en áreas del periodismo inmersas en constante innovación (García-Avilés *et al.*, 2018) y citan a Bleyen *et al.*, 2014 para destacar la necesidad de estudiar y analizar la innovación periodística desde la perspectiva de los profesionales en tanto que esta actividad no siempre aparece reflejada en los trabajos realizados por la Academia (Bleyen *et al.*, 2014 citado en García-Avilés *et al.*, 2018). Asimismo, el corpus de estudios precedentes que usan esta aproximación desde antes de la explosión de la IA generativa (Noain-Sánchez, 2022) y que llega a la investigación reciente (Cazzamatta y Sarisakaloğlu, 2025; Cuartielles *et al.*, 2024; Fieiras-Ceide *et al.*, 2025; Gutiérrez-Caneda *et al.* 2024) subrayan su pertinencia. Se realizaron 16 entrevistas a dos grupos: 1) investigadores académicos y expertos en el uso de las herramientas de IA en tareas de verificación; 2) profesionales de la verificación tanto de organizaciones independientes como de áreas de verificación de los medios. Las entrevistas, cuya duración oscila entre los 30 y los 70 minutos, se llevaron a cabo entre enero de 2025 y junio de 2025, tanto de manera presencial como a través de la herramienta de videoconferencia Teams, si así lo requería la heterogeneidad geográfica de los participantes, y siempre de manera sincrónica.

Todas las entrevistas fueron grabadas y, posteriormente, transcritas para su análisis. En lo que respecta al tratamiento de los datos personales de los entrevistados, el proceso se realizó conforme al RGPD (DOUE, 2016) y la LOPDGDD (BOE, 2018). No se anonimizó el nombre y apellidos de los entrevistados en tanto que sus declaraciones se justifican en referencia al cargo público al que representan (Tabla 1). Asimismo, el estudio no comenzó hasta obtener el visto bueno del Comité de Ética de la Investigación de la URJC (registro interno: 120120250322025).

Para la transcripción se aplicó, como primer paso, la proporcionada automáticamente por Teams, aunque todas fueron revisadas de manera manual por dos personas para evitar errores. Posteriormente, se realizó un análisis textual cualitativo mediante Atlas.ti, *software* que propicia una codificación inductivo-je-rárquica y permite asignar categorías y relaciones para identificar temas recurrentes. Si bien este abordaje en tanto que cualitativo no permite una generalización estadística, como se ha indicado sí aporta información valiosa sobre las observaciones y experiencias de los profesionales de la verificación y académicos (Cazzamatta *et al.*, 2025).

3.1. Selección de la muestra

Las entrevistas semiestructuradas se realizaron a un total de 16 informantes divididos en dos grupos: Grupo 1) conformado por 11 profesionales del *fact-checking*, provenientes tanto de organizaciones independientes, como de áreas de verificación ligadas a medios de comunicación, tanto de titularidad pública o privada. Todos ellos debían cumplir los siguientes criterios: a) pertenecer a una entidad española, b) cuyo idioma fuese preminentemente el castellano y c) registrada en las bases de datos internacionales de verificadores de referencia (Duke Reporters' Lab, signataria del código de principios de la IFCN o bien, EDMO). Grupo 2) consta de 5 académicos que investigan las herramientas de IA y su uso en las tareas de verificación (Tabla 1).

El Grupo 1 integra a personas con un papel destacado en entidades de verificación e incluye desde profesionales del periodismo, hasta ingenieros y desarrolladores vinculados a departamentos de innovación, por lo que el guion de estas entrevistas hubo de adaptarse a los distintos perfiles, generando dos versiones (perfil de comunicación y perfil técnico). Ahora bien, dado que una de las consecuencias de la inclusión de la IA en el quehacer periodístico es la aparición de perfiles híbridos (Cano-Orón *et al.*, 2024, p. 6; Rojas-Torrijos *et al.*, 2023; Sixto-García *et al.*, 2021) encontramos perfiles mixtos de manera recurrente tal y como se indica en la Tabla 1. A este respecto, es pertinente señalar que no fue posible realizar las entrevistas previstas con EFE y Verificat, lo que constituye una limitación relevante del estudio, dado el papel que desempeñan en el ecosistema de la verificación. Asimismo, la sobrerrepresentación de RTVE se explica por el hallazgo, durante la primera entrevista realizada con Verifica RTVE, de la existencia de varias unidades operativas con funciones de verificación desarrolladas de manera independiente dentro de la propia organización.

Tabla 1. Perfil de los entrevistados.

Entrevistado	Perfil	Institución	Ocupación	Modalidad	Fecha
Javier Cantón (JC)	Académico/profesional	Universidad de Granada	Investigador	remoto	12/03/2025
Iris López de Solís (ILS)	Profesional/académico	RTVE-Archivo	Responsable de procesos de IA en archivo	presencial	02/06/2025
Uxía Carral (UC)	Académico/profesional	Universidad Carlos III	Investigador	remoto	12/05/2025

Entrevistado	Perfil	Institución	Ocupación	Modalidad	Fecha
Patricia Ruiz (PR)	Profesional	<i>Maldita.es</i>	Responsable entidad de verificación	presencial	04/03/2025
Roger Cuartielles (RC)	Académico/profesional	Universitat Pompeu Fabra	Investigador	remoto	12/05/2025
David Corral (DC)	Profesional	RTVE	Responsable de innovación	remoto	12/03/2025
Sergio Arce (SA)	Académico	Universidad de La Rioja	Investigador	remoto	08/06/2025
Carmen Pérez Cernuda (CPC)	Profesional	RTVE	Responsable proyectos de IA	presencial	02/06/2025
Manuel Álvarez Rufs (MAR)	Académico	UNED	Investigador	remoto	27/05/2025
Borja Díaz-Merry (BD)	Profesional	Verifica RTVE	Responsable área de verificación	presencial	25/02/2025
Rubén Míguez (RM)	Profesional	<i>Newtral</i>	Responsable de procesos de IA	remoto	23/04/2025
Virginia Bazán (VB)	Profesional	RTVE-Archivo	Responsable de archivo	presencial	02/06/2025
Pablo Hernández (PH)	Profesional/académico	<i>Maldita.es</i>	Coordinador de investigación académica	presencial	04/03/2025
Carla Pina (CP)	Profesional	Infovéritas	Responsable entidad de verificación	presencial	02/04/2025
Tomás Rudich (TR)	Profesional	<i>Newtral</i>	Coordinador del equipo de verificación	remoto	26/03/2025
Iván López Olmos (ILO)	Profesional	TVE	Periodista	presencial	29/04/2025

Fuente: elaboración propia.

4. Resultados

4.1. Caracterización de las herramientas de IA en tareas de verificación

De acuerdo con las respuestas de los informadores, existe un corpus heterogéneo de herramientas de IA cuya primera clasificación atiende a dos ejes concretos: 1) quién diseña la herramienta y 2) fase del proceso de verificación en que se usa.

4.1.1. Según la procedencia de la herramienta de IA

Respecto al primer eje, los entrevistados mencionan, por una parte, los desarrollos comerciales, esto es, softwares que ya están en el mercado, de uso generalizado y que conllevan «sus pros y sus contras, porque la misma empresa que extiende la mentira es la que da herramienta para decirte que es mentira» (DC). El desconocimiento de la programación de los algoritmos que integran estas herramientas añade una primera limitación, ya que «los verificadores deben ceñirse a lo que la propia matriz tecnológica permita» (MAR) y suscita reticencias: «ni periodistas, ni *fact-checkers* tenemos capacidad real de acceso, ni de control de los algoritmos que están en manos de unas cuantas corporaciones tecnológicas» (MAR).

La lista de herramientas gratuitas usadas por los verificadores es más amplia. Entre estas, se mencionan aquellas financiadas por instituciones «usamos los últimos desarrollos de Vera.ai, que es la evolución del proyecto WeVerify de la Unión Europea y su InVID-WeVerify verification plug-in» (BD). A partir de

aquí, se describen una serie de desarrollos propios, es decir, creados ad hoc, bien en solitario, bien mediante consorcios con universidades e institutos de investigación que parten, igualmente, «de los grandes modelos de razonamiento fundacionales como Gemini de Google, Open AI, DeepSeek... de las grandes tecnológicas» (RM) pero adaptados al proceso propio de verificación de cada *fact-checker*.

Según los entrevistados, tanto las implementaciones gratuitas, como las desarrolladas exprofeso resultan más útiles para el trabajo de los verificadores que las aplicaciones comerciales. Por ejemplo, el proyecto IVERES, desarrollado por RTVE junto con varias universidades públicas y compartido con EFE Verifica, se menciona como una herramienta modular de gran utilidad «para seguimiento de redes sociales, verificación de vídeos ultra falsos e investigación de audios falsos» (BD). Pero, además, sirve de apoyo en otras tareas, como traducir voces o texto en lenguas minoritarias: «Si hablan, por ejemplo, en urdu, en farsi, en idiomas que es complicado que ninguno de nosotros hablemos, nos dan una traducción, al menos en inglés y, si es posible, en español» (DC) y lo hacen con gran eficacia: «Trint, que es de pago por suscripción, hace algo de transcripción multilingüe, pero IVERES da un abanico de más de 200 idiomas y el nivel de eficacia y fiabilidad es muy alto» (BD). Estas herramientas parten, además, de una visión común entre desarrolladores y periodistas, como en el caso de ClaimDetection y ClaimHunter, creadas por el equipo de ingenieros de *Newtral* en colaboración con los verificadores de la redac-

ción. Por último, se destaca la importancia del software libre y «de las comunidades de programadores de que se dedican a construir el código y que este sea compartido» (MAR).

4.1.2. Según la fase en la que se aplican las herramientas de IA

El catálogo de herramientas con IA usadas por los *fact-checkers* incluye todas aquellas que se aplican en tres de los momentos básicos del proceso, esto es: 1) en la fase previa, de prospección o de detección de potenciales bulos, 2) en la propia verificación y 3) en la posterior difusión de resultados. Su capacidad para automatizar tareas permite ahorrar tiempo en cualquiera de los tres estadios mencionados. Además, se emplean como soporte en tareas complementarias propias de rutinas periodísticas, tales como la síntesis informativa, la traducción multilingüe y la transcripción automática de textos. Los entrevistados subrayan su potencial en la fase de detección. En esta, se emplean algoritmos capaces de rastrear patrones virales, recolección de afirmaciones (*claims*), identificar narrativas emergentes y anticipar la propagación de bulos. De este modo, posibilitan una intervención preventiva frente a la desinformación, al centrarse en el «*prebunking* o detección temprana, en lugar de trabajar el *debunking*» (JC). Esto se traduce en un mayor volumen de contenidos analizados en menos tiempo, en una fase, la de detección, que «se lleva una buena parte del tiempo de los verificadores» (JC). Esta identificación precoz propicia encontrar patrones comunes y elementos capaces de conformar narrativas y estudiar su evolución «en un panel de narrativas» (PH) a la vez que se sigue la progresión de los bulos individuales. «Si no se tiene una narrativa para ese potencial bulo, contenido desinformador o lo que sea, con IA se sugiere una nueva narrativa» (PR). En última instancia posibilita conocer si estos contenidos ya han sido investigados por *fact-checkers* de otros países y por tanto constituyen una adaptación en otro idioma y/o contexto, al buscar en «una base de datos global».

Antes no podíamos saber si un verificador externo había hecho una información o no» (RM). Por otra parte, los chatbots creados por los propios verificadores constituyen recursos altamente útiles y ambivalentes, ya que tienen eficacia tanto en la fase de detección, como en la de distribución. En la detección, se consolidan como un canal a través del cual reciben consultas de los ciudadanos, alertan a los *fact-checkers* y alimentan «una base de datos súper nutrida y muy potente de todos esos contenidos que pueden ser desinformaciones» (PR). En la difusión, si la cuestión recogida por el *chatbot* ya ha sido comprobada y existe un dictamen sobre su veracidad, «el sistema empareja la consulta con la verificación ya realizada y permite difundirlo inmediatamente» (RC). De nuevo, al ser multilingüe, además de buscar comprobaciones en bases de datos de todo el mundo, «puede responder en distintos idiomas» (TR). Se potencia así la propagación de los contenidos ya investigados y comprobados, «otro de los grandes retos a los que se enfrenta la verificación, pues el verificado debe llegar tan lejos como el bulo o, por lo menos, a aquellas personas a las que les ha llegado» (JC).

Asimismo, el uso de la IA en esta fase proporciona información sobre «cuál es el momento óptimo para lanzar un contenido» (TR). En la fase de investigación la IA ofrece niveles de eficacia desigual. Hay varios proyectos en desarrollo, «incluso herramientas de IA para desmentir a la IA. Pero son proyectos experimentales, que de momento no se pueden utilizar, ni siquiera con supervisión humana, porque no están suficientemente alimentadas como para poder fiarse de los resultados» (UC). Por el momento, estas herramientas constituyen únicamente «indicios que sirven para tener una pista, pero no para estar 100 % seguros» (TR). Sin embargo, su uso no es desdeñable: «cualquier motor de búsqueda está impulsado por IA y una de las técnicas principales en verificación, la búsqueda inversa, está apoyada en IA» (BD). En este aspecto y dado que la desinformación realizada con IA es «más visual y auditiva» (TR), los entrevistados mencionan prioritariamente el uso de plataformas para comprobar vídeos, imágenes o audios. La IA es capaz de encontrar pistas que no percibe el humano, como «tramas que no son correctas, que no corresponden a lo que debería ser un vídeo normal» (DC) y, en el caso de los audios, la capacidad de contrastar declaraciones con las voces reales de personalidades clonadas ayuda a conocer si una voz ha sido creada por IA y «hacer ataques inversos» (DC). Permiten, además, un análisis profundo de contenidos dudosos pues perciben «las pequeñas huellas digitales que puede dejar la IA [...] y realizan una definición de todos los objetos que hay [en una imagen] lo que a un humano le costaría muchísimo» (ILO).

4.2. Integración, autonomía y supervisión de las herramientas de IA

4.2.1. Nivel de eficacia y fiabilidad del *fact-checking* automatizado

En lo referido a la aplicación de herramientas computacionales para evaluar la veracidad de un contenido potencialmente falso mediante algoritmos que lo analizan en profundidad, lo contrastan con bases de datos, corpus documentales y fuentes previamente verificadas y emiten un juicio sobre su grado de veracidad sin intervención humana directa (García-Marín, 2022) la postura de los verificadores no es muy optimista. En sus primeras formulaciones, se asumía que la IA podría operar de manera autónoma en la comprobación de cualquier tipo de contenido y ofrecer resultados con altos niveles de fiabilidad. Bajo esta premisa, el periodista introduciría un *input* (el contenido dudoso) y la herramienta lo clasificaba automáticamente como verdadero o falso y generaba un *output* (el correspondiente desmentido). Al respecto, los testimonios recogidos evidencian una unanimidad en torno a la imposibilidad de automatizar por completo la verificación: «no se trata de volcar un vídeo en una herramienta y que te resuelva las dudas, sino que es un trabajo al final humano, periodístico. Las herramientas nos apoyan, pero no deciden por nosotros» (TR). De hecho, es en el *fact-checking* automatizado donde menos efectividad presenta la IA: son «sistemas probabilísticos que una vez aciertan y otra vez no» (PR) son poco fiables y dependiendo del fotograma que aparece en el vídeo o, incluso, del tipo de herramienta que utilice, dan resultados muy distintos» (CP).

Actualmente, «no hay una herramienta que tenga la autonomía para hacer todo el proceso, el *fact-checking* automatizado es como buscar el Santo Grial» (RC). Las limitaciones son de diversos tipos. En primer lugar, no existen, como en otros campos, grandes volúmenes de datos para el entrenamiento de los sistemas de IA y que esta sea fiable. Por ejemplo, «en la detección facial o la detección de caracteres ópticos, hay muchos *dataset* estandarizados y aceptados a nivel mundial, pero en verificación no hay un corpus para entrenar estos modelos, ni el inglés, ni en español, en ningún idioma» (JC). Si bien existen proyectos de *fact-checking* automatizado experimentales por todo el mundo, «todavía no han configurado un gran *dataset* que permita hacer un entrenamiento a gran escala y en el idioma de cada uno» (JC). Por otra parte, tiene que ser un sistema multimodal, que tenga en cuenta infinidad de parámetros, «no solo el lenguaje sino el propio contenido de las imágenes, es decir, un *dataset* completo donde haya mucho etiquetado por parte de humanos para que el entrenamiento sea correcto y que tenga sus frutos» (JC).

La rapidez indisoluble del ejercicio del periodismo constituye otro de los inconvenientes: las herramientas de IA necesitan entrenarse con datos de actualidad, pero «los desmentidos corren tan rápido, que la máquina todavía no tiene el contexto que necesita para desmentirlos» (UC). En cuanto a la codificación se debe tener en cuenta que «los algoritmos reducen la complejidad social y hacen simplificaciones y abstracciones» (MAR) y que las herramientas poseen limitaciones heredadas de la dificultad para conceptualizar ideas o de la falta de consenso a la hora de definir ciertos términos: «si no somos capaces de ponernos de acuerdo en qué es un discurso de odio, por ejemplo, va a ser difícil programar una herramienta de detección de discurso de odio» (MAR). Las herramientas de IA se conciben como indicadores: «nos dan indicios, pero siempre las usamos con mucha prudencia» (TR) es decir, funcionan como «un avisador que nos dice fíjate en esto, que tiene probabilidad de ser falso» (SA), lo que hace es «subrayar aquello que no cuadra con la historia» (ILO), pero «como verificador absoluto, no» (RC).

4.2.2. Nivel de autonomía y supervisión humana

Dado que la IA se plantea no como sustituto del juicio profesional, sino como instrumento complementario o conjunto de estrategias «a las que se puede recurrir en cada paso de la comprobación, aplicándolos por capas» (RM) se prioriza el análisis periodístico del redactor. La necesidad de supervisión humana es clara: «aunque las herramientas tengan una conclusión muy clara, siempre hay un margen de error que es importante [...] el ojo humano todavía ve muchas cosas que la inteligencia todavía no puede ver» (BD). Tras usarlas para encontrar indicios, se realiza una investigación «tradicional, periodística [...] llamar a las fuentes pertinentes y preguntarles si algo ha sucedido o no, etcétera» (TR). Además, no todas las verificaciones son susceptibles de ser automatizadas: «no me imagino un *bot* llamando a la Policía Nacional y que esta les desmienta algo. Esa parte del proceso no se puede automatizar, será un

híbrido» (RM). El enfoque que se plantea corresponde al denominado *human in the loop*, situando al «ser humano en el centro y lo que hay son una serie de herramientas que utilizan automatizaciones para cada fase del proceso de trabajo de verificación» (JC).

4.3. Implicaciones éticas

4.3.1. Veracidad de la información: fuentes, datos fiables y sesgos

Hay coincidencia en señalar que los desafíos éticos tradicionalmente vinculados al ejercicio del periodismo no han cambiado, aunque sí se han incrementado algunos de ellos y destacan, particularmente, los referidos a la veracidad y exactitud de los datos. Esta merma proviene, en primer lugar, de la mencionada dificultad para disponer de *datasets* adecuados para entrenar a los algoritmos, a lo que se suma la complejidad para trasladar a su programación la complejidad sociocultural, elementos propios del lenguaje y principios de la profesión periodística. Esto provoca que el trabajo con IA demande una «continua supervisión por falta de contexto» (UC) y que esta sea realizada por el profesional adecuado: «no nos olvidemos de que los datos tienen sesgos porque nosotros como sujetos los tenemos» (UC). En segundo lugar, se subraya la complejidad para enseñar a los modelos de IA a discriminar entre fuentes fiables y aquellas que no lo son, puesto que «empiezan a acumular datos y se pierden. ChatGPT, Perplexity, por ejemplo, están diseñadas para obtener información de todos los sitios, de todo lo que hay en la red» (SA). Este reto puede acrecentarse porque hasta ahora, «la IA se estaba entrenando con contenidos hechos por humanos, pero, de ahora en adelante, va a tomar muchos elementos que ha generado la propia IA» (CP).

A esto se le añade que «puede que llegue un punto en que, cuando busquemos algo en vías como Internet, aparezca más contenido desinformativo que real e informativo» (UC). Incluso hay «campañas que inundan Internet, para que ChatGPT o Perplexity encuentren datos falsos y los tomen como verdaderos» (SA). Se trata del *grooming de los lenguajes* o LLM *grooming*, una técnica «destinada a entrenar estos lenguajes con desinformación dirigida» (ILS) para, por ejemplo, manipular las respuestas de los *chatbots* y propagar desinformación. En esta «nueva cadena informativa alterada por la IA» (PR) el consumo de información se da en un entorno «marcado por la *enshittification*, en un buscador, tú buscas algo y, de diez imágenes que aparecen, tres están generadas con IA. Eso también es desinformación y es alterar nuestro flujo de información» (PR). Una solución sería crear «una IA específica que bebiera solo de fuentes fiables para verificar» (SA) o bien, «tratar de desarrollar sistemas públicos algorítmicos» (MAR). La posibilidad de que campañas de desinformación condicionen los modelos de IA son factores cuya connivencia compromete la fiabilidad del proceso. Es precisamente en este punto donde los profesionales vinculados a archivo reivindican su labor, «el trabajo de documentación es en sí mismo un trabajo de verificación [...] el gestor documental, donde almacenamos los contenidos es información verificada, fiable y validada con criterios humano» (VB). «Si tuviéramos que prescindir de la IA, sería malo, pero

si tuviéramos que prescindir del archivo, sería una tragedia» (CPC).

4.3.2. Uso de la IA y transparencia

Hay acuerdo en subrayar la importancia de informar desde el principio sobre si se ha aplicado la IA, en qué parte del proceso se ha aplicado y cuáles son las herramientas usadas, «ser transparentes, eso es lo más importante de todo» (UC). En la misma línea: «Siempre que se usa algún tipo de IA se indica al final del artículo, en las fuentes y herramientas que se han utilizado para el artículo» (CP). Esta declaración «debe quedar muy clara, con un *disclaimer* que diga esto puede contener errores, porque lo he hecho una máquina» (RM) para que «cualquier persona que lo esté leyendo, sepa que las herramientas no son infalibles» (CP). Asimismo, una de las razones por las que se enfatiza que el trabajo ha sido realizado por un ser humano es porque la percepción por parte del lector de una aplicación intensiva de la IA redundaría en una pérdida de credibilidad y confianza: «A todos nos gusta un poco de innovación, pero puede llegar a un punto de valle donde nos empiece a asustar tanta IA» (UC).

4.4. Códigos éticos y de deontología profesional y autorregulación

Aunque se están empezando a desarrollar códigos de uso de la IA, por el momento, a excepción de una norma interna mencionada por los profesionales de RTVE, los informantes no refieren normas escritas en las entrevistas: «las tengo muy claras internamente, pero no lo hemos puesto por escrito» (RM). «Cree-mos que son útiles, pero no tenemos código por escrito por una cuestión de tiempo» (TR). A este respecto, algunos de los verificadores se remiten a estándares externos: «Formamos parte de la IFCN que tiene una serie de guías, directrices sobre cómo debemos trabajar los *fact-checkers*, la metodología, la transparencia, las fuentes, la rigurosidad. Cualquier protocolo de IA en la verificación tiene que estar ajustado a estos lineamientos» (TR). Se advierte que estos estándares y directrices externos son declaraciones laxas y su utilidad es limitada. En este sentido, se demanda el desarrollo de guías, recomendaciones y códigos éticos más precisos que, si bien deben ser «congruentes con las organizaciones paraguas de las que forman parte los *fact-checkers*» (TR) deben adaptarse a las necesidades de la profesión periodística. Su valía depende de que sean consensuados por todos los actores involucrados: «Sería fundamental juntarnos todos para ver el uso que le podemos dar o hasta dónde puedes llegar a trabajar con la IA» (CP), más aún, cuando en España no existe un colegio de periodismo capaz de ejercer labores de control.

La principal razón que se esgrime para la inexistencia de códigos éticos escritos proviene de la complejidad que conlleva el desarrollo. En primer lugar, hay que «diferenciar la ética del proceso en el que se construye una herramienta de IA y, por otro lado, la ética de la estructura o sistema social al que va a afectar esa IA como agente cognitivo y social» (MAR). Es decir, conjugar los principios del sistema algorítmico con los usos y normas socioculturales. A esto, se le suma la necesidad de incluir en el diseño

algorítmico los principios del sector concreto en el que se aplica: «Tanto si es periodismo o *fact-checking*, son fundamentales los valores de transparencia, de objetividad, neutralidad, imparcialidad» (MAR). En la misma línea, deben: «Incluir todo lo ligado a la ética periodística: Verificar con hechos que sean totalmente contrastados, no basarse en rumores y no estar haciendo elucubraciones, ir ligado a lo que debe de ser un buen periodista» (SA). A pesar de la casi ausencia de normas éticas escritas, se constata que los verificadores realizan prácticas de autorregulación: «No usamos la IA para redactar artículos, sino para apoyarnos en la labor de verificar» (CP) o bien: «No se usa para generar imágenes de alguien que tengan sus derechos de autor, no tomamos la imagen de ningún famoso y hacemos que hable sin su consentimiento. Todo lo que no haríamos sin la IA, no lo hacemos con la IA» (ILO). Finalmente, también se concibe como instrumento necesario el desarrollo de un marco jurídico más concreto que el actual Reglamento de IA que clarificase su uso en labores periodísticas.

5. Discusión y conclusiones

Aplicada en la verificación de contenidos, la IA permite optimizar las acciones preliminares (monitorización, búsqueda inversa o análisis previo) así como las destinadas a la distribución de contenidos ya verificados, tal y como describen Canavilhas e Ito (2025) por el contrario, la fase central de verificación, aunque sirve de apoyo, sigue requiriendo intervención humana. El *fact-checking* completamente automatizado presenta limitaciones estructurales que impiden su autonomía plena. En consonancia con lo observado por Noble (2018) y O'Neil (2016), para quienes su bajo nivel de fiabilidad, la falta de neutralidad de las herramientas y la opacidad en el desarrollo de los algoritmos propician que sea imposible sustituir el juicio humano.

Entre las limitaciones estructurales que presentan las herramientas actuales y que condicionan su desempeño, se encuentran, entre otras: a) La inexistencia de conjuntos de datos robustos, específicos y multilingües para entrenar modelos orientados a la verificación impide alcanzar estándares aceptables de automatización, en línea con lo adelantado por García-Marín, 2022. b) La falta de contextualización semántica y sociocultural: los algoritmos carecen de capacidad para interpretar matices lingüísticos y culturales, lo que genera errores en narrativas complejas y reduce la precisión en entornos informativos heterogéneos. c) La variabilidad de resultados: herramientas como detectores de vídeo y audio ofrecen únicamente indicios, no conclusiones definitivas, lo que obliga a contrastar manualmente cada hallazgo.

Estas limitaciones tienen implicaciones prácticas relevantes. En primer lugar, la IA debe aplicarse «por capas», actuando como un sistema de alertas que señala posibles falsedades sin determinar conclusiones. En segundo lugar, la supervisión humana resulta indispensable para validar hallazgos, contactar fuentes y garantizar el rigor periodístico. Finalmente, la percepción social condiciona su integración: un uso intensivo puede erosionar la credibilidad del verificador, como también advierten estudios previos (Pérez-Seijo *et al.*, 2020).

La tendencia en la integración de la IA en los procesos de verificación apunta hacia herramientas híbridas que combinan IA con bases de datos verificadas, evitando el reentrenamiento constante y reforzando la trazabilidad del proceso. Pero, sobre todo, se destaca un uso instrumental y complementario, bajo el enfoque *human-in-the-loop* (Milosavljević y Vobič, 2019) que sitúa al periodista en el centro del proceso, otorgándole la responsabilidad de la toma de decisiones y relegando la IA a un papel de apoyo en tareas específicas. Esta «línea en la arena» Johnson (2024, p. 1) frente al proceso automatizado permite a los periodistas retomar el control respecto los criterios de publicación y, por ende, la autoridad para rendir cuentas. «El uso de técnicas *human-in-the-loop* garantiza que cualquier contenido manipulado con IA haya pasado antes por la revisión de un profesional responsable antes de su publicación» (Quiles-Morán y Arias-Robles, 2024).

En cuanto al aspecto ético, es reseñable que, como principal desafío, se señale la falta veracidad y exactitud de los datos y resultados obtenidos. No sólo porque emerge en contraposición al principio periodístico considerado como eje normativo del ejercicio profesional: «el compromiso ético fundamental del periodista es con la verdad» (FAPE, 2017) sino por la naturaleza de la labor concreta que se analiza, la verificación. Ello se acentúa por las limitaciones inherentes a los *datasets*, opacidad y sesgos algorítmicos, así como por la falta de neutralidad en los sistemas, lo que concuerda con los planteamientos de Ventura-Salom y Gelado-Marcos (2025) y la dependencia de las grandes tecnológicas (Simon, 2022). La merma en la veracidad de los datos y los sesgos algorítmicos identificada en este trabajo amplían las conclusiones de Ufarte-Ruiz *et al.* (2021) quienes ya advertían que la IA no elimina los riesgos tradicionales del periodismo, tampoco crea otros nuevos, pero sí los amplifica.

La transparencia se indica como otro de los requerimientos. Se reseña la importancia de informar al usuario sobre el uso de IA, lo que apuntala las recomendaciones del Instituto Poynter (2024) y se alinea con lo expuesto por Ufarte-Ruiz *et al.* (2021) y el principio de explicabilidad propuesto por Cortina (2019). Según Ufarte-Ruiz *et al.*, «los lectores tienen derecho a entender cómo se emplea la IA y las decisiones que se adoptan en términos comprensibles» (2021, p. 67), lo que refuerza la pertinencia de incorporar principios que garanticen trazabilidad algorítmica, incluyendo información sobre quién desarrolla los sistemas, con qué sesgos y con qué objetivos (Cortina, 2019). El «derecho a entender» significa que esta información debe presentarse en términos accesibles que permitan comprender las decisiones derivadas del procesamiento de datos (Hansen *et al.*, 2017).

Sanahuja-Sanahuja *et al.* (2022, 2025) añaden a la transparencia, la supervisión humana como pilares fundamentales en la aplicación de la IA en procesos periodísticos. Igualmente, el Consell de la Informació de Catalunya identifica la transparencia como

uno de los ocho principios esenciales, dando a conocer a los usuarios «la existencia de los algoritmos y las características básicas de su funcionamiento» (Ventura, 2021:35). Aunque el Consell reconoce que no todos los procesos algorítmicos se podrán explicar, «por razones industriales o simplemente por la complejidad de sus mecanismos» sí es necesario que los usuarios conozcan el grado de automatización «de modo que, si hallan en él inexactitudes, puedan saber a qué atribuir las» (Ventura, 2021:35). Aunque la transparencia se encuentra recogida en los códigos de la IFCN y la EFCSN, su implementación específica en el contexto español sigue siendo limitada (Cuartielles *et al.*, 2024). Junto a las iniciativas voluntarias, se necesitan avances normativos que regulen la visibilidad de los procesos algorítmicos en el ámbito periodístico, sin embargo, en España persiste la ausencia de mecanismos que informen al consumidor cuando una noticia ha sido generada mediante sistemas automatizados (Túñez-López *et al.*, 2021).

En cuanto al desarrollo de códigos éticos, se constata la casi ausencia de guías internas. Los participantes abogan por: a) una regulación consensuada a nivel sectorial, esto es, desde el mundo del periodismo, b) que cuente con todos los agentes que intervienen en la integración de la IA en el periodismo c) y que sea compartida a nivel global, al estilo de la cooperación internacional que propone la Carta de París sobre IA y Periodismo (RSF, 2023). Esta regulación ha de integrar los principios del periodismo tradicional con las exigencias derivadas de las tecnologías emergentes y ha de permitir adaptar las guías a las particularidades de cada medio. De su flexibilidad dependerá su eficacia, en una coyuntura, que impele a adoptar una filosofía de mutación constante en los medios (Fieiras-Ceide *et al.* 2023). La necesidad de consensuar principios éticos se contempla en paralelo a la urgencia de crear un marco jurídico sectorial, esto es, centrado en el ejercicio del periodismo y/o agentes afines, pero con vocación internacional. Dicho marco, permitiría dirimir responsabilidades y derechos de los actores implicados al detallar desde la propiedad de los modelos de IA y el diseño y uso de algoritmos, hasta la responsabilidad sobre los resultados generados.

Cabe reseñar que ninguno de los entrevistados ha mencionado la posibilidad de crear un colegio oficial de periodismo mecanismos de gran utilidad para la autorregulación de la actividad (Real-Rodríguez, 2018) y que actuarían como refuerzo de la labor de los profesionales ante los desarrollos de la IA (Flores-Vivar, 2025).

6. Financiación y apoyos

Esta publicación se inscribe dentro del proyecto «Desaf_IA: Desafíos, usos y limitaciones de la IA en el fact-checking y la lucha contra la desinformación» (2024/SOLCON-135623), financiado por la Universidad Rey Juan Carlos dentro del programa Proyectos Impulso 2024, entre el 1 de enero 2025 y el 31 de mayo de 2026

7. Contribución de autores

Conceptualización	Ideas; formulación o evolución de los objetivos y metas generales de la investigación.	Autor 1
Curación de datos	Actividades de gestión para anotar (producir metadatos), depurar datos y mantener los datos de la investigación (incluido el código de software, cuando sea necesario para interpretar los propios datos) para su uso inicial y su posterior reutilización.	Autores 1,2 y 3
Análisis formal	Aplicación de técnicas estadísticas, matemáticas, computacionales u otras técnicas formales para analizar o sintetizar datos de estudio.	Autores 1,2 y 3
Adquisición de fondos	Adquisición del apoyo financiero para el proyecto que conduce a esta publicación.	Autores 1,2 y 3
Investigación	Realización de una investigación y proceso de investigación, realizando específicamente los experimentos, o la recolección de datos/evidencia.	Autor 1,2 y 3
Metodología	Desarrollo o diseño de la metodología; creación de modelos.	Autor 1
Administración del proyecto	Responsabilidad de gestión y coordinación de la planificación y ejecución de la actividad de investigación.	Autor 1
Recursos	Suministro de materiales de estudio, reactivos, materiales, pacientes, muestras de laboratorio, animales, instrumentación, recursos informáticos u otras herramientas de análisis.	Autores 1, 2 y 3
Software	Programación, desarrollo de software; diseño de programas informáticos; implementación del código informático y de los algoritmos de apoyo; prueba de los componentes de código existentes.	Autor 1
Supervisión	Responsabilidad de supervisión y liderazgo en la planificación y ejecución de actividades de investigación, incluyendo la tutoría externa al equipo central.	Autor 1
Validación	Verificación, ya sea como parte de la actividad o por separado, de la replicabilidad/reproducción general de los resultados/experimentos y otros productos de la investigación.	Autores 1 y 2
Visualización	Preparación, creación y/o presentación del trabajo publicado, específicamente la visualización/presentación de datos.	Autor 3
Redacción / Borrador original	Preparación, creación y/o presentación del trabajo publicado, específicamente la redacción del borrador inicial (incluyendo la traducción sustantiva).	Autor 1
Redacción / Revisión y edición	Preparación, creación y/o presentación del trabajo publicado por los miembros del grupo de investigación original, específicamente revisión crítica, comentario o revisión, incluidas las etapas previas o posteriores a la publicación.	Autores 1, 2 y 3

8. Declaración sobre uso de inteligencia artificial

En este artículo se ha utilizado la inteligencia artificial para realizar la primera transcripción de las entrevistas. La herramienta a las que se ha recurrido es la integrada en la aplicación Teams. Gracias a ese uso de la inteligencia artificial se ha conseguido una primera transcripción de las entrevistas, que posteriormente hubo de ser revisada.

9. Referencias bibliográficas

Alonso-González, M. y Sánchez Gonzales, H. (2024). Inteligencia artificial en la verificación de la información política. Herramientas y tipología. *Más Poder Local*, (56), 27-45. <https://doi.org/10.56151/maspoderlocal.215>

Beckett, C. y Yaseen, M. (2023). Generating Change. A global survey of what news organisations are doing with AI. *The London School of Economics*

and Political Science, Google News Initiative. <https://www.journalism.ai/info/research/2023-generating-change>

Bleyen, V. A., Lindmark, S., Ranaivoson, H. y Ballon, P. (2014): A typology of media innovations: Insights from an exploratory study”. *The Journal of Media Innovations*, 1(1), 28-51.

Canavilhas, J. e Ito, L. (2025). On Fact-Checking Service: Artificial Intelligence’s Uses in Ibero-American Fact-Checkers. *Social Sciences*, 14(9), 514. <https://doi.org/10.3390/socsci14090514>

Cano-Orón, L. y López-Meri, A. (2024). *Introducción al uso de la IA en periodismo. Guía de referencias y modos de uso.* Universitat de València.

Cazzamatta, R. y Sarisakaloğlu, A. (2025). AI-Generated Misinformation: A Case Study on Emerging Trends in Fact-Checking Practices Across Brazil, Germany, and the United Kingdom. *Emerging Media*, 3(2), 214251. <https://doi.org/10.1177/27523543251344971>

- Comisión Europea. (2025). Código de Buenas Prácticas para la IA de uso general. <https://digital-strategy.ec.europa.eu/es/politicas/contents-code-gpai>
- Cortina, A. (2019). *Ética de la inteligencia artificial. Anales de la Real Academia de Ciencias Morales y Políticas*, 96 (2018-2019), 379-394. Boletín Oficial del Estado. https://www.boe.es/biblioteca_juridica/anuarios_derecho/abrir_pdf.php?id=ANU-M-2019-10037900394.
- Costello, T. H., Pennycook, G. y Rand, D. G. (2024). Durably reducing conspiracy beliefs through dialogues with AI. *Science*, 385(6714). <https://doi.org/10.1126/science.adq1814>
- Cuartielles, R., Mauri-Ríos, M. y Rodríguez-Martínez, R. (2024). Transparencia en el uso de la IA en las plataformas de fact-checking en España y sus desafíos éticos, *Communication & Society*, 37(4), 257-271. <https://doi.org/10.15581/003.37.4.257-271>.
- Federación de Asociaciones de Periodistas de España, FAPE. (2017). *Código Deontológico*. <https://fape.es/home/codigo-deontologico-1/>
- EFCSN. (2022). *European Code of Standards for Independent Fact-Checking Organisations*. <https://efcsn.com/code-of-standards/>
- Fieiras Ceide, C. F., Tüñez López, M. y Fernández Lombao, T. (2025). De la frecuencia al algoritmo: implementación de IA en las emisoras de radio de España. *Revista Latina De Comunicación Social*, (83), 1-22. <https://doi.org/10.4185/rlds-2025-2457>
- Flores-Vivar, J. M. (9 de septiembre de 2025) En la era de la IA (y de la desinformación), el ejercicio del periodismo debe ser refrendado por un colegio de periodistas. *Cultura de Red* (Blogs Fundación para el Conocimiento Madrid+d). <https://www.madrimasd.org/blogs/culturadered/en-la-era-de-la-ia-y-de-la-desinformacion-el-ejercicio-del-periodismo-debe-ser-refrendado-por-un-colegio-de-periodistas/>
- Floridi, L. (2019). *The Logic of Information: A Theory of Philosophy as Conceptual Design*. Oxford University Press.
- Full Fact (2025). *Full Fact AI*. <https://fullfact.org/ai/>
- García-Avilés, J. A., Carvajal Prieto, M. y Arias Robles, F. (2018). Implantación de la innovación en los cybermedios españoles: análisis de las percepciones de los periodistas. *Revista Latina De Comunicación Social*, (73), 369-384. <https://doi.org/10.4185/RLCS-2018-1260>
- García-Marín D. (2022). Modelos algorítmicos y fact-checking automatizado. Revisión sistemática de la literatura. *Documentación de las Ciencias de la Información*, 45(1), 7-16. <https://doi.org/10.5209/dcin.77472>
- García-Marín, D. (2024). Periodismo contra la desinformación. Proceso y estructura de las verificaciones en el fact-checking. *Infonomy*, 2(2). <https://doi.org/10.3145/infonomy.24.026>.
- García-Marín, D. (2025). Verificación digital y fact-checking. Teoría, método y herramientas. En García-Marín, D. (Ed.), *Desinformación y fact-checking en la era de la IA*. (pp. 95-125). CIESPAL.
- García-Marín, D., Rubio-Jordán, V. y Salvat-Martinrey, G. (2023). Chequeando al fact-checker. Prácticas de verificación política y sesgos partidistas en Newtral (España). *Revista de Comunicación*, 22(2), 207-223. <https://doi.org/10.26441/RC22.2-2023-3184>
- Gelado-Marcos, R., Ventura-Salom, B., Poch-Butler, S. L. y De-la-Calle-Velasco, G. (2025). ¿Cazando bulos o desenmarañando narrativas desinformadoras? El fact-checking ibérico y la lucha contra la desinformación. *Revista Mediterránea de Comunicación/Mediterranean Journal of Communication*, 16(2), e28237. <https://www.doi.org/10.14198/MEDCOM.28237>
- Graves, L. (2018). Boundaries Not Drawn: Mapping the institutional roots of the global fact-checking movement. *Journalism Studies*, 19(5) 613-631.
- Guo, Z., Schlichtkrull, M. y Vlachos, A. (2022). A survey on automated fact-checking. *Transactions of the association for computational linguistics*, 10, 178-206. https://doi.org/10.1162/tacl_a_00454
- Gutiérrez-Caneda B., Lindén, C-G. y Vázquez-Herrero, J. (2024) Ethics and journalistic challenges in the age of artificial intelligence: talking with professionals and experts. *Front. Commun.* 9:1465178. <https://doi.org/10.3389/fcomm.2024.1465178>
- Hansen, M., Roca-Sales, M., Keegan, J. M. y King, G. (2017). *Artificial intelligence: Practice and implications for journalism*. Tow Center for Digital Journalism, Columbia University. <https://doi.org/10.7916/D8X92PRD>
- Huwyler, L (22 de diciembre de 2022) *Investiga con la Caja de herramientas avanzada de VerificaRTVE*. <https://www.rtve.es/noticias/20221221/investiga-verificartve-caja-herramientas-avanzada/2412478.shtml>
- EDMO (July 25, 2025). *Israel-Iran War: For the first time, AI content takes the lead in disinformation*. <https://edmo.eu/wp-content/uploads/2025/07/EDMO-Horizontal-.pdf>
- IFCN. (2024) *IFCN Code of Principles*. <https://ifcnco-deofprinciples.poynter.org/>
- IPSOS (22 de julio de 2025). *Percepciones globales contradictorias sobre la IA: señales mixtas para las marcas*. <https://url-shortener.me/1GPI>
- Johnson, P. R. (2024). A Case of Claims and Facts: Automated Fact-Checking the Future of Journalism's Authority. *Digital Journalism*, 12(10), 1461-1484. <https://doi.org/10.1080/21670811.2023.2174564>
- Quiles-Morán, B. y Arias-Robles, F. (2024). Verificación automática bajo vigilancia humana: implementación de IA en el caso Newtral. *RAE-IC, Revista de la Asociación Española de Investigación de la Comunicación*, 11 (Especial). <https://doi.org/10.24137/raeic.11.e.3>
- La-Barbera, D., Roitero, K. y Mizzaro, S. (2022). A hybrid human-in-the-loop framework for fact checking. In *Proceedings of the 6th Workshop on natural language for artificial intelligence (NL4AI 2022)*, 3287. <https://ceur-ws.org/Vol-3287/paper4.pdf>
- Larraz, I., Míguez, R. y Sallicati, F. (2023). Semantic similarity models for automated fact-checking: ClaimCheck as a claim matching tool. *Profesional de la Información*, 32(3). <https://doi.org/10.3145/epi.2023.may.21>
- Ley Orgánica 3/2018, de 5 de diciembre, de Protección de Datos Personales y garantía de dere-

- chos digitales (LOPDGDD). <https://www.boe.es/buscar/act.php?id=BOE-A-2018-16673>
- Milosavljević, M. y Vobič, I. (2019). Human Still in the Loop: Editors Reconsider the Ideals of Professional Journalism Through Automation. *Digital Journalism*, 7(8), 1098-1116. <https://doi.org/10.1080/21670811.2019.1601576>
- Moreno Gil, V. (2024). Aplicación de la inteligencia artificial en las plataformas de fact-checking españolas. En Parrat Fernández, S., Mayoral Sánchez, J. y Chaparro Domínguez, M. Á. (eds.) *Periodismo e inteligencia artificial. Aplicaciones y desafíos profesionales*. Comunicación Social Ediciones y Publicaciones. <https://doi.org/10.52495/c2.emcs.25.p108>
- Moreno-Gil, V. y Salgado-de Dios, F. (2023). El cumplimiento del código de principios de la International Fact-checking Network en las plataformas de verificación españolas. Un análisis cualitativo. *Revista De Comunicación*, 22(1), 293-307. <https://doi.org/10.26441/RC22.1-2023-2971>
- Noain-Sánchez, A. (2022). Addressing the impact of artificial intelligence on Journalism: the perception of experts, journalists, and academics. *Communication & society*, 35(3), 105-121. <https://doi.org/10.15581/003.35.3.105-121>
- Noble, S. U. (2018). *Algorithms of Oppression: How Search Engines Reinforce Racism*. NYU Press.
- O'Neil, C. (2016). *Weapons of Math Destruction: How Big Data Increases Inequality and Threatens Democracy*. Crown Publishing.
- Pérez-Seijo, S., Gutiérrez-Caneda, B. y López-García, X. (2020). Periodismo digital y alta tecnología: de la consolidación a los renovados desafíos. *Index.comunicacion*, 10(3), 129-151. <https://doi.org/10.33732/ixc/10/03period>
- Poynter (2024). *AI ethics guidelines. A starter kit for newsrooms*. <https://www.poynter.org/ai-ethics-journalism/ai-ethics-guidelines/>
- Real-Rodríguez, E. (2018). La profesión periodística ante sus retos éticos: autorregulación profesional y comunicativa frente a regulación. La situación en España. *Estudios sobre el Mensaje Periodístico*, 24(1), 341-360. <http://dx.doi.org/10.5209/ESMP.59954>
- Reglamento (UE) 2016/679 del Parlamento Europeo y del Consejo, de 27 de abril de 2016, (Reglamento general de protección de datos, RGPD). <https://www.boe.es/doue/2016/119/L00001-00088.pdf>
- Reglamento (UE) 2024/1689 del Parlamento Europeo y del Consejo, de 13 de junio de 2024, por el que se establecen normas armonizadas en materia de inteligencia artificial. <https://www.boe.es/doue/2024/1689/L00001-00144.pdf>
- Rojas-Torrijos, J. L. y Herrera-Damas, S. (2023). *Manual de nuevos perfiles profesionales para el periodismo y la no ficción*. Tirant lo Blanch.
- RSF (2023) *Paris Charter on AI and Journalism*. <https://url-shortener.me/1GLP>
- Sanahuja-Sanahuja, R. y López-Rabadán, P. (2025). Ethical guidelines for journalistic use of GenAI. The main trends in the international debate and progress in self-regulation in Spain *Communication & Society*, 38(1), 214-23. <https://doi.org/10.15581/003.38.1.016>
- Sanahuja-Sanahuja, R. y López-Rabadán, P. (2022). Ética y uso periodístico de la inteligencia artificial. Los medios públicos y las plataformas de verificación como precursores de la rendición de cuentas en España. *Estudios sobre el Mensaje Periodístico*, 28(4), 959-970. <https://dx.doi.org/10.5209/esmp.82385>
- Sánchez-González, M., Sánchez-Gonzales, H. M. y Martínez-Gonzalo, S. (2022). Inteligencia artificial en verificadores hispanos de la red IFCN: proyectos innovadores y percepción de expertos y profesionales. *Estudios sobre el Mensaje Periodístico*, 28(4), 867-879. <https://doi.org/10.5209/esmp.82735>
- Sanguinetti, P. (2023) *Tecnohumanismo*. LHG.
- Simon, F. M. (2022). Uneasy Bedfellows: AI in the News, Platform Companies and the Issue of Journalistic Autonomy. *Digital Journalism*, 10(10), 1832-1854. <https://doi.org/10.1080/21670811.2022.2063150>
- Singer, J. B. (2021). Border patrol: the rise and role of fact-checkers and their challenge to journalists' normative boundaries. *Journalism*, 22(8), 1929-1946. <https://doi.org/10.1177/1464884920933137>
- Sixto-García, J., Rodríguez-Vázquez, A. I. y López-García, X. (2021). Sistemas de verificación en medios nativos digitales e implicación de la audiencia en la lucha contra la desinformación en el modelo ibérico. *Revista de Comunicación de la SEECI*, 54, 41-61.
- Tejedor, S. y Sancho-Ligorred, B. (2023). Cartografía mundial de herramientas, fact-checkers y proyectos contra la infodemia. *Estudios sobre el Mensaje Periodístico*, 29(4), 933-942. <https://dx.doi.org/10.5209/esmp.87838>
- Torres-Toukourmidis, A., Lagares-Díez, N. y Barredo-Ibáñez, D. (2021). Accountability journalism during the emergence of COVID-19: Evaluation of transparency in official fact-checking platforms. En Rocha, Á., Reis, J.L., Peter, M.K., Cayolla, R., Loureiro, S., Bogdanović, Z. (Eds) *Marketing and Smart Technologies*. Smart Innovation, Systems and Technologies, vol 205. Springer (pp. 561-572). https://doi.org/10.1007/978-981-33-4183-8_44
- Túñez-López, J. M., Fieiras Ceide, C. y Vaz-Álvarez, M. (2021). Impacto de la Inteligencia Artificial en el Periodismo: transformaciones en la empresa, los productos, los contenidos y el perfil profesional. *Communication & Society*, 34(1), 177-193. <https://www.doi.org/10.15581/003.34.1.177-193>
- Ufarte Ruiz, M. J., Calvo-Rubio, L. y Murcia-Verdú, F. (2021). Los desafíos éticos del periodismo en la era de la inteligencia artificial. *Estudios sobre el Mensaje Periodístico*, 27(2), 673-684. <https://doi.org/10.5209/esmp.69708>
- Ventura, P. (2021). Algoritmos en las redacciones: Retos y recomendaciones para dotar a la inteligencia artificial de los valores éticos del periodismo. Consell de la Informació de Catalunya. https://cic.periodistes.cat/wp-content/uploads/2022/03/algorismes_a_les_redaccions_ESP.pdf
- Ventura-Salom, B. y Gelado Marcos, R. (2025). Vigilando al vigilante: análisis de los sesgos ideológicos en los contenidos analizados por las agencias verificadoras de referencia en España y Portugal. *Comunicación y Sociedad*, 1-26. <https://www.comunicacionsociedad.cucsh.udg.mx/index.php?journal=comsoc&page=article&op=view&path%5B%5D=e8900&path%5B%5D=6820>

- Vizoso, A. y Vázquez-Herrero, J. (2019). Plataformas de *factchecking* en español. Características, organización y método. *Communication & Society*, 32(1), 127-144.
- Vosoughi, S., Roy, D. y Aral, S. (2018). The spread of true and false news online. *Science*, 359(6380), 1146-1151. <https://doi.org/10.1126/science.aap9559>

Amaya Noain-Sánchez. PDI en el Departamento de Ciencias de la Comunicación y Sociología de la Universidad Rey Juan Carlos de Madrid. Doctora Europea, Universidad Complutense de Madrid - Goethe Universität Frankfurt am Main. Su tesis doctoral obtuvo el Premio AEPD 2015- Categoría Investigación y cuenta con otros reconocimientos como Emerald Literati Outstanding Paper, 2017. Licenciada en Periodismo, UCM y Máster en Radio, UCM-RNE. Ha formado parte de varios proyectos de investigación en España y Alemania, y codirige el Proyecto «Desaf_IA: Desafíos, usos y limitaciones de la IA en el fact-checking y la lucha contra la desinformación» (2024/SOLCON-135623). ORCID: <https://orcid.org/0000-0002-2845-3285>

Leticia Rodríguez-Fernández. Profesora Titular en el Departamento de Marketing y Comunicación de la Universidad de Cádiz. Doctora por la Universidad Complutense de Madrid. Directora General de Comunicación Estratégica de la Universidad de Cádiz. Vocal de la Conferencia de Rectores y Rectoras de las Universidades Españolas (CRUE) en el Foro contra las Campañas de Desinformación en el ámbito de la Seguridad Nacional. Investigadora Principal del Grupo de Investigación Comunicación estratégica e influencia en la sociedad digital e investigadora en el Proyecto «Desaf_IA: Desafíos, usos y limitaciones de la IA en el fact-checking y la lucha contra la desinformación» (2024/SOLCON-135623). ORCID: <https://orcid.org/0000-0002-7472-5472>

Marina Santín-Durán. Profesora Titular en el Departamento de Ciencias de la Comunicación y Sociología de la Universidad Rey Juan Carlos de Madrid. Secretaria Académica de la Facultad de Ciencias de la Comunicación. Doctora en Ciencias de la Información por la Universidad Complutense de Madrid. Licenciada en Periodismo y Derecho por la misma universidad. Forma parte de varios grupos de investigación y ha dirigido y participado en numerosos proyectos de investigación tanto nacionales como europeos. Investigadora en el Proyecto «Desaf_IA: Desafíos, usos y limitaciones de la IA en el fact-checking y la lucha contra la desinformación» (2024/SOLCON-135623). ORCID: <https://orcid.org/0000-0002-0036-0861>