bid

Revista científica editada por la Facultad de Información y Medios Audiovisuales de la Universidad de Barcelona
ISSN 1575-5886

BiD 55 (diciembre 2025)
DOI: 10.1344/bid2025.55.07

ARTÍCULO ORIGINAL

Principales editores científicos en los canales de Telegram: una aproximación a la detección de *fake channels* con ChatGPT y DeepSeek

Major Academic Publishers on Telegram Channels: An Approach to Fake Channel Detection Using ChatGPT and DeepSeek

Principals editors científics als canals de Telegram: una aproximació a la detecció de canals falsos amb ChatGPT i DeepSeek

Víctor Herrero-Solana
Universidad de Granada
victorhs@ugr.es

Carlos Castro-Castro
Universidad de Granada
ccastro@ugr.es ✉

Recibido: 10-09-2025
Aceptado: 07-11-2025

Citación recomendada

Herrero-Solana, Víctor, & Carlos Castro-Castro (2025). Principales editores científicos en los canales de Telegram: una aproximación a la detección de *fake channels* con ChatGPT y DeepSeek. BiD, 55. https://doi.org/10.1344/BID2025.55.07

Resumen

Objetivos: Identificar la existencia de canales falsos en Telegram que suplantan a grandes editoriales académicas, evaluar la efectividad de los Modelos de Lenguaje a Gran Escala (LLMs), específicamente ChatGPT y DeepSeek, para su detección, analizar las fuentes web utilizadas por estos modelos y determinar la posible existencia de sesgos geográficos en dichas fuentes.

Metodología: Selección de 13 grandes editoriales académicas del portal SCImago: Elsevier, Springer, Wiley-Blackwell, Routledge, OUP, de Gruyter, Brill, CUP, IEEE, Hindawi, World Scientific, Nature y Thieme. Identificación de 37 canales de Telegram asociados y aplicación de un prompt estandarizado a ChatGPT y DeepSeek para evaluar la autenticidad de cada canal, con la función de búsqueda web activada. Análisis comparativo de las respuestas de los modelos y verificación mediante clasificación manual.

Resultados: Se identificó que el 78,38% de los canales analizados eran fraudulentos. Ambos modelos mostraron una alta efectividad en la detección de canales falsos, pero limitaciones significativas para validar canales legítimos. Se observaron diferencias metodológicas: DeepSeek adoptó un enfoque contextual, mientras que ChatGPT requirió una verificación explícita. El análisis de las fuentes web reveló que DeepSeek priorizó contenido de ciberseguridad, mientras que ChatGPT utilizó predominantemente fuentes institucionales y redes sociales, con un claro predominio de fuentes occidentales en ambos casos.

Palabras clave

Telegram; canales falsos; editores académicos; LLM; ChatGPT; DeepSeek, detección de desinformación; verificación de fuentes.

Abstract

Objectives: To identify the existence of fake channels on Telegram that impersonate major academic publishers, evaluate the effectiveness of Large Language Models (LLMs), specifically ChatGPT and DeepSeek, for their detection, analyze the web sources used by these models, and determine the potential existence of geographical biases in these sources.

Methodology: Selection of 13 major academic publishers from the SCImago portal: Elsevier, Springer, Wiley-Blackwell, Routledge, OUP, de Gruyter, Brill, CUP, IEEE, Hindawi, World Scientific, Nature, and Thieme. Identification of 37 associated Telegram channels, and application of a standardized prompt to ChatGPT and DeepSeek to evaluate the authenticity of each channel, with the web search function enabled. Comparative analysis of the models' responses and verification through manual classification.

Results: It was identified that 78.38% of the analyzed channels were fraudulent. Both models showed high effectiveness in detecting fake channels but significant limitations in validating legitimate channels. Methodological differences were observed: DeepSeek adopted a contextual approach, while ChatGPT required explicit verification. The analysis of web sources revealed that DeepSeek prioritized cybersecurity content, whereas ChatGPT predominantly used institutional sources and social media, with a clear predominance of Western sources in both cases.

Keywords

Telegram; fake channels; academic publishers, LLM, ChatGPT;, DeepSeek; misinformation detection; source verification.

Resum

Objectius: Identificar l'existència de canals falsos a Telegram que suplanten les principals editorials acadèmiques, avaluar l'eficàcia dels models grans del llenguatge (LLM), ChatGPT i DeepSeek, per a la seva detecció, analitzar les fonts web utilitzades per aquests models i determinar la possible existència de biaixos geogràfics en aquestes fonts.

Metodologia: Selecció de 13 grans editorials acadèmiques a partir del portal SCImago: Elsevier, Springer, Wiley-Blackwell, Routledge, OUP, de Gruyter, Brill, CUP, IEEE, Hindawi, World Scientific, Nature i Thieme. Identificació de 37 canals de Telegram associables i aplicació d'un prompt estandarditzat a ChatGPT i DeepSeek per avaluar l'autenticitat de cada canal, activant la funció de cerca web. Anàlisi comparativa de les respostes dels models i verificació mitjançant classificació manual.

Resultats: Es va identificar que el 78,38% dels canals analitzats eren fraudulents. Ambdós models van mostrar alta efectivitat en la detecció de canals falsos, però limitacions significatives en la validació de canals reals. Es van observar diferències metodològiques: DeepSeek va adoptar un criteri contextual, mentre que ChatGPT va requerir verificació explícita. L'anàlisi de fonts web va revelar que DeepSeek va prioritzar continguts de ciberseguretat, mentre que ChatGPT va utilitzar predominantment fonts institucionals i xarxes socials, amb clar predomini de fonts occidentals en ambdós casos.

Paraules clau

Telegram, canals falsos, editors acadèmics, LLM, ChatGPT, DeepSeek, detecció de desinformació; verificació de fonts.

1. Introducción

La aplicación de mensajería instantánea Telegram ha superado hace tiempo su función inicial de mensajería entre personas, convirtiéndose en una tecnología sólida para el intercambio global de información. A diferencia de las redes sociales clásicas, los usuarios suelen asociarla con comunicación íntima (Dargahi et al., 2017), pasando desapercibido que los canales de este servicio pueden constituir una fuente sustancial de información. La falta de control centralizado de publicación de canales hace fácil que el usuario pueda verse confundido con el origen y responsabilidad de estos canales, tal como ya se ha demostrado para el caso de la prensa española (Herrero y Castro, 2022).

La proliferación de noticias falsas y desinformación se ha convertido en una preocupación significativa en la era digital. Con el auge de las plataformas de redes sociales y la creciente facilidad de difusión de información, la propagación de contenido engañoso plantea serias amenazas a la opinión pública, los procesos democráticos y la armonía social (Shu et al. 2017). En este contexto, los modelos grandes del lenguaje (LLM) han emergido como herramientas prometedoras para combatir la desinformación, gracias a su profundo conocimiento del mundo y sus potentes capacidades de razonamiento. Investigaciones recientes han demostrado que modelos preentrenados ajustados, tanto modelos antiguos como BERT y RoBERTa o los LLM más avanzados, pueden alcanzar resultados notables en la detección de noticias falsas sin necesidad de datos auxiliares extensos (Flores-Vivar y García-Peñalvo, 2023). Estos modelos han mostrado particular eficacia mediante estrategias de *prompting*, especialmente el razonamiento en cadena de pensamiento (*chain-of-thought*), que mejora significativamente el rendimiento en tareas de verificación de hechos (Chen y Shu, 2024).

Para el presente estudio, se seleccionaron dos LLM representativos del panorama actual: ChatGPT y DeepSeek. La elección de ChatGPT se debe a ser el primero en irrumpir (y llamar mucho la atención) a finales de 2022 (Thorp, 2023) y que actualmente tiene aún una posición dominante en el mercado de modelos de lenguaje. Según diversos análisis de mercado, ChatGPT mantiene una cuota superior al 80% entre los *chatbots* de IA generativa, consolidándose como el LLM más utilizado tanto por usuarios individuales como por empresas (StatCounter 2025). Su amplia adopción y reconocimiento lo convierten en el referente natural para evaluar capacidades de detección de contenido fraudulento.

Por su parte, la inclusión de DeepSeek obedece a su singular relevancia en el momento de comenzar el diseño de este experimento. En enero de 2025, esta *startup* china irrumpió en los medios internacionales al presentar su modelo R1, que demostró capacidades comparables a las de sus competidores occidentales, pero desarrollado bajo restricciones significativas de acceso a hardware avanzado. Las políticas de control de exportaciones de Estados Unidos, implementadas desde octubre de 2022, restringieron severamente a fabricantes como Nvidia la venta de sus chips más potentes (H100 y A100) a empresas chinas. DeepSeek logró entrenar sus modelos utilizando chips H800, una versión con rendimiento limitado diseñada específicamente para el mercado chino. En lugar de limitarse al uso estándar del *framework* de Nvidia (CUDA), sus ingenieros recurrieron a programación de muy bajo nivel en PTX —el lenguaje tipo ensamblador sobre el que se apoya CUDA— para exprimir al máximo el potencial de estos chips (Chen, 2025). El impacto de su aparición generó incluso una fuerte caída bursátil de la propia Nvidia: casi 600 *billions* en un día (Carew et al., 2025). Este caso representa un interesante contrapunto metodológico frente a ChatGPT, al tratarse de un modelo desarrollado bajo restricciones tecnológicas que, sin embargo, alcanza niveles de rendimiento competitivos.

2. Marco teórico

2.1. Los canales de Telegram

El éxito de Whatsapp como alternativa a los SMS cuando se comenzó a generalizar el uso de teléfonos móviles con pantalla táctil conectados a Internet, marcó el inicio del crecimiento de las aplicaciones de mensajería instantánea. La aparición de WeChat en China y aplicaciones como Telegram y Signal, presentadas como herramientas más transparentes y seguras, hizo nacer un entorno comunicativo nuevo. Los diferentes avances de las distintas aplicaciones han ido enriqueciendo mutuamente las capacidades de cada una de ellas, lo que ha contribuido a un aumento constante de usuarios (Gregorio et al., 2017).

La incorporación de los canales en Telegram en 2015 supuso el surgimiento de una nueva forma de comunicación unidireccional que permite difundir noticias, contenido de entretenimiento, comunicaciones oficiales, etc., a audiencias globales ilimitadas (Kitsa, 2023). La incorporación de novedades desde entonces ha sido una constante, como la edición de mensajes (que permite corregir errores o actualizar información después de la publicación), herramientas básicas de análisis para rastrear el rendimiento de los mensajes, mejora de las capacidades de gestión y diversificación formatos y herramientas. Telegram no ha sido una aplicación estándar ya que algunas de sus decisiones sobre criptografía y seguridad son singulares (Jakobsen, 2015). Su arquitectura de seguridad se presenta como un valor primordial para la solidez de la propia herramienta (Wardle y Derakhshan, 2017).

Telegram ha sido incorporado a la actividad informativa en medios de comunicación (Sánchez y Martos, 2020), y se ha comprobado la solidez de los canales para la difusión informativa (Sedano y Palomo, 2018). Telegram se ha hecho un hueco en el panorama mediático actual al establecerse como un canal alternativo para el consumo de noticias. Se ha demostrado que la confianza en la marca periodística juega un papel crucial en la selección de las fuentes de información en esta plataforma (Martos y Sánchez, 2024). Su flexibilidad ha permitido desarrollar soluciones informativas y comunicativas en situaciones tan extraordinarias, como la guerra de Ucrania, donde se han llegado a crear nuevos formatos comunicativos (Steblyna, 2024). Utilizados también por instituciones universitarias de España y Latinoamérica, para su uso en difusión de su información (Cisternas et al., 2022). No ha ocurrido en la misma medida en entornos puramente comerciales. Telegram se ha mostrado sólido para soportar sistemas de información documental y, al mismo tiempo, generar entornos comunicativos a diferentes niveles. Estas capacidades se han comprobado en diversos formatos (Mohammed et al., 2024). La variedad de los mismos, la solidez de su sistema de almacenamiento y la capacidad difusión de contenidos se han usado con éxito en experiencias de aprendizaje de idiomas y reforzamiento de capacidades lingüísticas (Abu-Ayfah, 2020).

Telegram proporciona un modelo estructurado para su uso en la interacción social y la moderación experta (Ghaffari et al., 2017). Dichas fortalezas no han dado los resultados deseados en campos como el sanitarios. En 2020, con motivo de la pandemia, desde Telegram se realizó un ofrecimiento para que las autoridades sanitarias de todo el mundo verificaran sus canales oficiales. Sorprendentemente, la iniciativa solo fue seguida por una veintena de países y no tuvo un gran recorrido. Paradójicamente, en esas circunstancias fue la herramienta elegida por muchos organismos para difundir su información en esta crisis (López, 2020). En esa misma época, ante la demanda del teletrabajo, Telegram mejoró notablemente las herramientas de videocomunicación, tanto en chats, como en grupos y canales, alcanzando capacidades incluso mayores que las de algunas de las más populares plataformas de streaming, desarrollándose como

un sistema híbrido que ha ampliado notablemente las capacidades comunicativas (Dargahi et al., 2021).

Los canales de Telegram progresivamente se han convertido en verdaderas plataformas de difusión (Willaert, 2023). El perfeccionamiento de los permisos avanzados para administradores los dota de capacidades para asignar roles y gestionar de forma granular quién puede publicar, editar, fijar o eliminar mensajes, convirtiéndolos funcionalmente en editores (Gregorio, 2020). La posibilidad añadida de asociar grupos de discusión a los canales unida a la mejora de las herramientas analíticas y estadísticas ha permitido además disponer de instrumentos para evaluar el alcance y la interacción de las publicaciones (La Morgia et al., 2023).

El sistema de búsqueda de Telegram mezcla, la localización de canales por las palabras contenidas en su título, con la localización de palabras en los contenidos de todos los canales suscritos. Resultando muy rápido y eficaz para la localización de contenidos en los canales suscritos, pero muy confuso e inseguro en la localización precisa de canales. (Jalilvand y Neshati, 2020).

En las actualizaciones de Telegram durante 2024 se ha reforzado de manera notable todo lo relativo a las mini apps en su plataforma de bots, con una creciente actividad alrededor de los premios y la monetización de servicios y actividades comerciales. El éxito de estos nuevos campos de actividad ha abierto una línea de actuación que, en las primeras actualizaciones de 2025, se está proyectando sobre los canales. El lanzamiento del sistema de verificación de terceros presentado por Telegram en la primera actualización del año es una buena prueba de ello. Este nuevo sistema para combatir estafas y desinformación une al tradicional check la posibilidad de que servicios oficiales ajenos al propio Telegram, verifiquen cuentas y chats. Los canales o cuentas verificados mediante este método podrán mostrar un icono único junto a su nombre, y al hacer clic, se podrá ver qué servicio realiza la certificación y el motivo. En la actualización de marzo de 2025, se han añadido utilidades para el control de la bandeja de entrada. Estas permiten establecer una tarifa para los mensajes entrantes de usuarios que no estén entre los contactos, lo que facilita un control total sobre la bandeja de entrada. Los usuarios podrán ganar estrellas (el sistema de pagos utilizado en las mini apps), filtrar los mensajes no deseados y evitar la sobrecarga de la bandeja de entrada. Esta configuración también se puede aplicar a los chats grupales y conversaciones de canales para mantener las interacciones enfocadas y libres de spam, lo que ayuda a sus propietarios a monetizar su comunidad y ganar estrellas a partir de conversaciones significativas (Telegram, 2025). Todas estas medidas permiten el aumentar la transparencia y la confianza en el contenido publicado, pero solo darán frutos, si los propios medios, las instituciones y la comunidad científica hace uso de estas.

Resulta indudable que estas utilidades han abierto oportunidades y nuevas vulnerabilidades, planteando retos a los que es necesario dar respuesta (Sušánka y Kokeš, 2017). El hecho de que cualquiera pueda crear un canal público de Telegram, imitar la imagen de un original, denominarlo e incluso alimentarlo con mensajes provenientes de cualquier fuente hace que el sistema no haya mostrado solidez, toda vez que la verificación de canales no se ha generalizado. Aunque existen multitud de canales "teóricamente oficiales" incluso enlazados desde sus webs, muchos aparecen en Telegram sin el check de verificación oficial (Herrero y Castro, 2022). En ocasiones, aparecen canales con denominaciones iguales, en los que resulta complicado verificar su naturaleza oficial, ya que incluso reproducen contenidos provenientes de los RSS de los propios medios y organizaciones (La Morgia et al., 2023). Hay evidencia de que las capacidades descritas, permiten la fácil difusión de desinformación debido a la privacidad y a la preservación del anonimato. Esto ha dado como resultado la proliferación de canales falsos, como se evidenció de manera notable tras la pandemia (Díez et al., 2021).

En los últimos años, tras varios episodios de resonancia mediática, parece que las relaciones con los organismos gubernamentales de diferentes países están siendo más fluidas y se están alcanzando más acuerdos con reguladores. Buena prueba de ello son las eliminaciones de millones de publicaciones y canales dañinos cada día, la publicación de informes diarios de transparencia y las líneas directas con ONGs para procesar las solicitudes de moderación con mayor rapidez (Durov, 2024).

La propia naturaleza de los canales, su facilidad de creación y la inexistente moderación (salvo la verificación o la eliminación respondiendo a una denuncia) hacen que la difusión de desinformación sea un hecho inevitable (Herasimenka et al., 2023). Su impacto es especialmente preocupante en el ámbito científico por los problemas de suplantación de identidad mediante el uso indebido de nombres de editoriales y logos oficiales para ganar credibilidad instantánea (Wardle y Derakhshan, 2017). La manipulación informativa mediante la difusión de estudios falsos, manipulados o incompletos puede influir negativamente en debates académicos y decisiones políticas (Cho et al., 2019). La inevitable erosión de la confianza, al confundir a los usuarios, puede dañar la credibilidad de la ciencia y las instituciones académicas (Allcott y Gentzkow, 2017).

La naturaleza de los canales de Telegram permite generar bases de información para analizarlas y obtener resultados fiables que pueden ser evaluados. Sin embargo, la labor de evaluación de los resultados de los análisis cuantitativos ha requerido hasta ahora el concurso de analistas humanos (La Morgia et al., 2018), los LLMs pueden abrir una nueva perspectiva en caso de poder prescindir de los humanos para esta tarea.

2.2. Irrupción de los LLMs: ChatGPT y Deepseek

El lanzamiento de ChatGPT por OpenAI en noviembre de 2022 provocó una convulsión que trascendió el ámbito de las TIC, al poner la inteligencia artificial generativa al alcance de personas sin conocimientos técnicos especializados. Millones de usuarios comenzaron a interactuar con la IA de manera sencilla, desencadenando una ola de cambios y debates que afectaron a la sociedad a nivel global. Empresas tecnológicas líderes en Silicon Valley reaccionaron rápidamente: Google lanzó Gemini, su modelo de lenguaje avanzado; Microsoft reforzó su apuesta con inversiones millonarias en OpenAI; Meta desarrolló sus propios modelos de lenguaje, como LLaMA y lo ofreció en abierto; X (anteriormente Twitter) presentó Grok, un modelo enfocado en la interacción social, entre otros muchos. Además, empresas como Amazon y Apple también intensificaron sus esfuerzos en IA, integrando tecnologías generativas en sus ecosistemas de servicios y dispositivos. Sin embargo, este auge se vio abruptamente interrumpido a principios de 2025 con la irrupción de DeepSeek, una startup china respaldada por el fondo de cobertura. DeepSeek, una alternativa sorprendentemente económica y de código abierto, no solo desafió el dominio de OpenAI y otras empresas occidentales. DeepSeek alteró casi todas las expectativas en el sector de la IA, demostrando que la innovación no está limitada a Silicon Valley y que el mercado global de inteligencia artificial es más dinámico y competitivo de lo que muchos anticipaban.

Más allá de las implicaciones socioeconómicas, ChatGPT y DeepSeek representan dos enfoques complementarios dentro de los LLMs, tecnologías que han redefinido el procesamiento del lenguaje natural mediante la arquitectura Transformer (Vaswani et al., 2017). Ambos sistemas operan mediante sofisticados mecanismos de atención, que asignan relevancia contextual a cada término, lo que permite la identificación de patrones y relaciones semánticas en grandes volúmenes de datos, que, combinada con sistemas de búsqueda en Internet, abre expectativas de análisis de información muy prometedoras.

ChatGPT se ha consolidado como un referente en la generación de diálogos fluidos y contextualizados. Su eficacia se basa en un proceso en dos fases: un preentrenamiento

masivo en corpus multilingües (libros, artículos, webs) y aprendizaje por refuerzo con retroalimentación humana (RLHF) (Ouyang et al., 2022), donde ajustadores humanos califican respuestas para priorizar coherencia y seguridad. Esta dualidad le permite adaptarse a tareas tan diversas como la redacción creativa o la simulación de asistencia técnica, con un riesgo inherente: la generación de alucinaciones (afirmaciones plausibles pero falsas) al extrapolar patrones de datos no verificados (Bender et al. ,2021).

DeepSeek es un LLM de código abierto optimizado para el análisis semántico profundo y la detección de desinformación. Aunque comparte la base técnica del Transformer, su diseño incorpora técnicas como Retrieval-Augmented Generation (RAG) (Lewis et al., 2020), que vincula el modelo a bases de datos externas, para contrastar afirmaciones en tiempo real. Además, su entrenamiento multilingüe especializado (español, chino, inglés) y su eficiencia computacional, lo posicionan como una herramienta escalable para procesar flujos masivos de datos.

La introducción de DeepSeek como una alternativa de código abierto y gratuita a ChatGPT, marca un punto de inflexión en el campo de la IA, ofreciendo nuevas posibilidades a la investigación científica. Si bien ambos modelos comparten fortalezas significativas en la mejora de la eficiencia y la democratización del acceso, las diferencias en su transparencia, costo y accesibilidad tienen implicaciones profundas para la innovación, la ética y la asignación de recursos. La transición de la IA de una herramienta asistencial, a un colaborador activo, requiere una adaptación continua de las prácticas de investigación (Kayaalp et al., 2025). Ambos modelos comparten la capacidad de realizar un análisis lingüístico profundo, identificando indicadores de manipulación emocional, ausencia de fuentes y contradicciones internas, elementos frecuentemente asociados a la difusión de informaciones falsas. (Shu et al., 2017).

2.3. Detección de desinformación con LLMs: avances y perspectivas

Los LLMs han transformado el campo de la detección de desinformación. Los modelos pre-entrenados como BERT y RoBERTa ya habían establecido una base sólida para la clasificación de textos (Liu et al., 2019). Sin embargo, los LLMs actuales, basados en arquitecturas Transformer, ofrecen capacidades superiores de comprensión contextual y razonamiento semántico, lo que permite abordar la desinformación de forma más adaptativa (Papageorgiou et al. 2025). Una contribución clave de los LLMs es su aplicación en escenarios de detección de noticias falsas con pocos ejemplos (*few-shot fake news detection).* Esta capacidad es crucial para adaptarse rápidamente a las nuevas narrativas de desinformación, ya que estos modelos pueden generalizar a partir de un número muy reducido de ejemplos, superando las limitaciones de los métodos supervisados tradicionales (Patel et al., 2025).

La integración de técnicas de explicabilidad (Explainable AI) se ha convertido en un pilar fundamental. Metodologías como LIME (Local Interpretable Model-agnostic Explanations) y SHAP (SHapley Additive exPlanations) permiten desglosar la base racional de una clasificación, identificando palabras o frases clave asociadas a contenido falso (Patel et al. 2025). Esto es esencial tanto para la auditoría del sistema como para proporcionar retroalimentación útil a los verificadores humanos. La investigación reciente también apunta hacia la detección multimodal, donde los LLMs se combinan con redes neuronales para procesar información de texto, imágenes y vídeos. La construcción de módulos de fusión de características permite identificar inconsistencias sutiles entre diferentes formatos, una característica común en las campañas de desinformación modernas (Lu y Yao, 2025).

Sin embargo, este potencial convive con desafíos importantes. Los LLMs son propensos a las alucinaciones –la generación de afirmaciones plausibles pero falsas– lo que

compromete su fiabilidad como detectores (Pooley, 2024). Asimismo, existe el riesgo de canibalismo del conocimiento: a medida que el contenido generado por IA inunda internet, los modelos futuros podrían entrenarse con salidas de otros modelos, degradando la calidad de su conocimiento base (Pooley, 2024). La detección de desinformación camuflada, que imita fuentes legítimas, exige que los modelos se centren en el análisis factual más que en patrones estilísticos superficiales (García, 2024).

En conjunto, los LLMs no son una solución única, sino herramientas poderosas cuya efectividad depende de una integración inteligente con métodos de verificación externa y una comprensión crítica de sus limitaciones. ChatGPT y DeepSeek ofrecen varias prometedoras capacidades, pero la que afecta al presente trabajo consiste en su capacidad de "leer" la web para procesar y "razonar" su contenido. De esta forma, el usuario puede crear rápidamente una herramienta de scraping sin necesidad de contar con conocimiento de codificación y beneficiándose del modelo de lenguaje para su análisis. A continuación, se detallan los pasos seguidos en este experimento.

3. Objetivos

El objetivo primario de este artículo abordaremos la presencia de los grandes editores científicos en estos canales y la posibilidad de detectar los posibles f*ake-channels* con ChatGPT y DeepSeek. Para ello no proponemos responder a las siguientes preguntas de investigación:

1. ¿Existen fake-channels en Telegram de los editores académicos más importantes?
2. ¿Podrían servir ChatGPT y Deepseek para identificar estos canales?
3. ¿Qué fuentes web utilizan estos LLMs para responder a la cuestión anterior?
4. ¿Existe algún sesgo en las fuentes por países en función del origen de cada uno de ellos?

4. Material y métodos

El presente estudio adopta un diseño de estudio de caso múltiple, siguiendo los lineamientos metodológicos establecidos por (Yin 2018). Según este autor, el estudio de caso constituye una investigación empírica que examina un fenómeno contemporáneo dentro de su contexto real, siendo especialmente apropiado cuando las fronteras entre el fenómeno y el contexto no son claramente evidentes. La elección de casos múltiples (13 editoriales y 37 canales asociados) responde a la lógica de replicación literal propuesta por Yin, donde cada caso sirve como una "réplica" que permite confirmar o refutar los patrones emergentes. Adicionalmente, como señala (Eisenhardt, 1989), la investigación basada en casos múltiples resulta particularmente adecuada para áreas temáticas novedosas —como en este caso— dado que permite generar teoría empíricamente fundamentada a partir de la comparación sistemática entre casos. A continuación, se presenta la justificación del caso de estudio atendiendo a sus antecedentes, el propósito de su selección y las unidades de análisis.

4.1. Antecedentes del caso

La proliferación de canales falsos en plataformas de mensajería constituye un fenómeno que se ha intensificado en los últimos años, especialmente tras la pandemia de CO-

VID-19. Telegram presenta características estructurales que lo hacen particularmente vulnerable a la suplantación de identidad. A diferencia de otras plataformas, Telegram se ha posicionado como un servicio con moderación mínima, lo que facilita la creación de canales que imitan a organizaciones legítimas (Urman y Katz, 2022). El impacto de estas vulnerabilidades en el ámbito de la comunicación científica resulta especialmente preocupante. Los escándalos de suplantación de identidad académica han sacudido el sector editorial en los últimos años, con casos documentados de autores falsos, cartas de aceptación fraudulentas y uso de inteligencia artificial para fabricar datos (Stockemer y Reidy, 2024). Esta erosión se extiende al ámbito científico cuando canales no verificados difunden información bajo la apariencia de editoriales académicas reconocidas.

Aunque Telegram ha implementado un sistema de verificación de terceros, su adopción por parte de las editoriales académicas ha sido prácticamente inexistente. Ninguna de las editoriales analizadas en este estudio contaba con canales verificados oficialmente en la plataforma. Esta ausencia de verificación oficial deja un vacío que los actores malintencionados explotan mediante el uso de nombres, logos y descripciones que imitan a las organizaciones legítimas.

4.2. Propósito de la selección

Criterios de selección de editoriales: Se seleccionaron las 13 principales editoriales académicas según el número de títulos de revista indexados en SCImago Journal y Country Rank (consulta enero 2025, N > 30k fuentes). El criterio de número de títulos, frente al de documentos publicados, se adoptó deliberadamente para evitar el sesgo que introducirían las *megajournals* y para garantizar una representación del ecosistema editorial en su diversidad. Las editoriales seleccionadas (Elsevier, Springer, Wiley-Blackwell, Routledge, Oxford University Press, De Gruyter, Brill, Cambridge University Press, IEEE, Hindawi, World Scientific, Nature y Thieme) representan los principales actores del sector y, por tanto, los blancos más probables de suplantación. El listado completo con las diferentes variantes se puede encontrar en el Apéndice I.

Justificación de la plataforma Telegram: La elección de Telegram como objeto de estudio responde a tres factores convergentes. En primer lugar, su creciente uso para la difusión de contenido académico y científico, documentado en estudios recientes sobre ecosistemas de desinformación (Pietro et al., 2024). En segundo lugar, sus limitaciones estructurales de verificación, que contrastan con plataformas más reguladas y generan vulnerabilidades específicas para la suplantación de identidad institucional (Herrero y Castro, 2022). En tercer lugar, la accesibilidad pública del contenido de los canales mediante URLs directas (t.me/s/[canal]), que permite su análisis mediante LLMs con capacidad de navegación web.

Justificación de los LLMs seleccionados: Ya lo habíamos adelantado en la introducción, la selección de ChatGPT y DeepSeek como herramientas de análisis responde a su representatividad dentro del panorama actual de modelos de lenguaje. ChatGPT, desarrollado por OpenAI, mantiene una posición dominante en el mercado occidental con una cuota superior al 80% entre los chatbots de IA generativa, consolidándose como el referente comercial para aplicaciones de verificación. DeepSeek representa un contrapunto metodológico relevante: un modelo de código abierto desarrollado bajo restricciones de acceso a hardware avanzado que, sin embargo, ha demostrado capacidades competitivas. La comparación entre ambos modelos permite evaluar si las diferencias en arquitectura, entrenamiento y orientación (comercial vs. código abierto / USA vs. China) influyen en la capacidad de detección de canales fraudulentos.

4.3. Unidades de análisis

Siguiendo la estructura jerárquica recomendada por (Yin 2018) para estudios de caso múltiple, este estudio define tres niveles de análisis:

Nivel	Unidad de análisis	N	Descripción
1	Editoriales académicas	13	Principales editoriales según SCImago por número de títulos
2	Canales de Telegram	37	Canales identificados mediante búsqueda por nombre de editorial
3	Evaluaciones de LLMs	74	ChatGPT (37) + DeepSeek (37)

Tabla 1. Unidades de análisis

Variables recopiladas para cada canal: nombre del canal, cuenta (@usuario), editorial a la que se asocia, número de seguidores, texto del perfil, clasificación emitida (REAL/FAKE/UNK) por cada LLM, fuentes web citadas por cada LLM, dominio geográfico de las fuentes, y finalmente etiquetado de referencia (*ground truth*) establecido por los investigadores

Delimitación temporal: Las evaluaciones se realizaron utilizando las versiones de ChatGPT-V4o y DeepSeek-V3, entre el 1 y el 15 de febrero de 2025, ambas con la función de búsqueda web activa.

Delimitación de alcance: Se incluyeron únicamente canales públicos cuyo nombre, descripción o contenido pudiera generar confusión con las editoriales oficiales. Se excluyeron grupos privados (no accesibles públicamente) y canales claramente paródicos o satíricos sin intención de suplantación.

4.4 Procedimiento de recopilación de datos

El procedimiento de recopilación siguió un protocolo estandarizado para garantizar la replicabilidad:

Identificación de canales: Para cada una de las 13 editoriales, se realizó una búsqueda en Telegram utilizando el nombre de la editorial. El sistema de búsqueda de Telegram muestra un máximo de 10 resultados por consulta. Se conservaron todos los canales que pudieran ser confundidos a simple vista con la editorial oficial.

Registro de datos: Para cada canal identificado se registró el nombre, la cuenta (@usuario), el número de seguidores y el texto del perfil en un dataset estructurado.

Evaluación mediante LLMs: Se lanzó a ChatGPT y DeepSeek el siguiente *prompt* estandarizado para cada canal:

> You are an expert in Telegram application channels. Could you indicate if the channel @[nombre_canal] on Telegram is an official channel of the publisher [editorial] or if it could be a fake channel? You can find the channel's content on the web https://t.me/s/[nombre_canal]. What facts do you base your assessment on to determine if it is fake or not?

El *prompt* se formuló en inglés para maximizar la capacidad de ambos modelos de acceder a fuentes web internacionales. Cada evaluación se realizó en un chat nuevo, sin activar el pensamiento profundo (*deep thinking*) y con la opción de búsqueda web activa.

Registro de respuestas: Se documentaron las respuestas completas de cada modelo, incluyendo la clasificación emitida y los dominios web citados como fuentes de verificación.

Verificación manual: Los investigadores establecieron una clasificación *ground truth* mediante verificación manual de cada canal, consultando las páginas web oficiales de las editoriales, sus perfiles verificados en otras redes sociales y el contenido histórico de los canales.

5. Resultados

En primer lugar, tenemos en la figura 1 un ranking de las editoriales con la composición de títulos de revistas por países. Encontramos un fuerte predominio de Estados Unidos, Reino Unido y Países Bajos, en general y de algunos países en editores concretos como Alemania en Springer, de Gruyter y Thieme o Egipto en Hindawi.

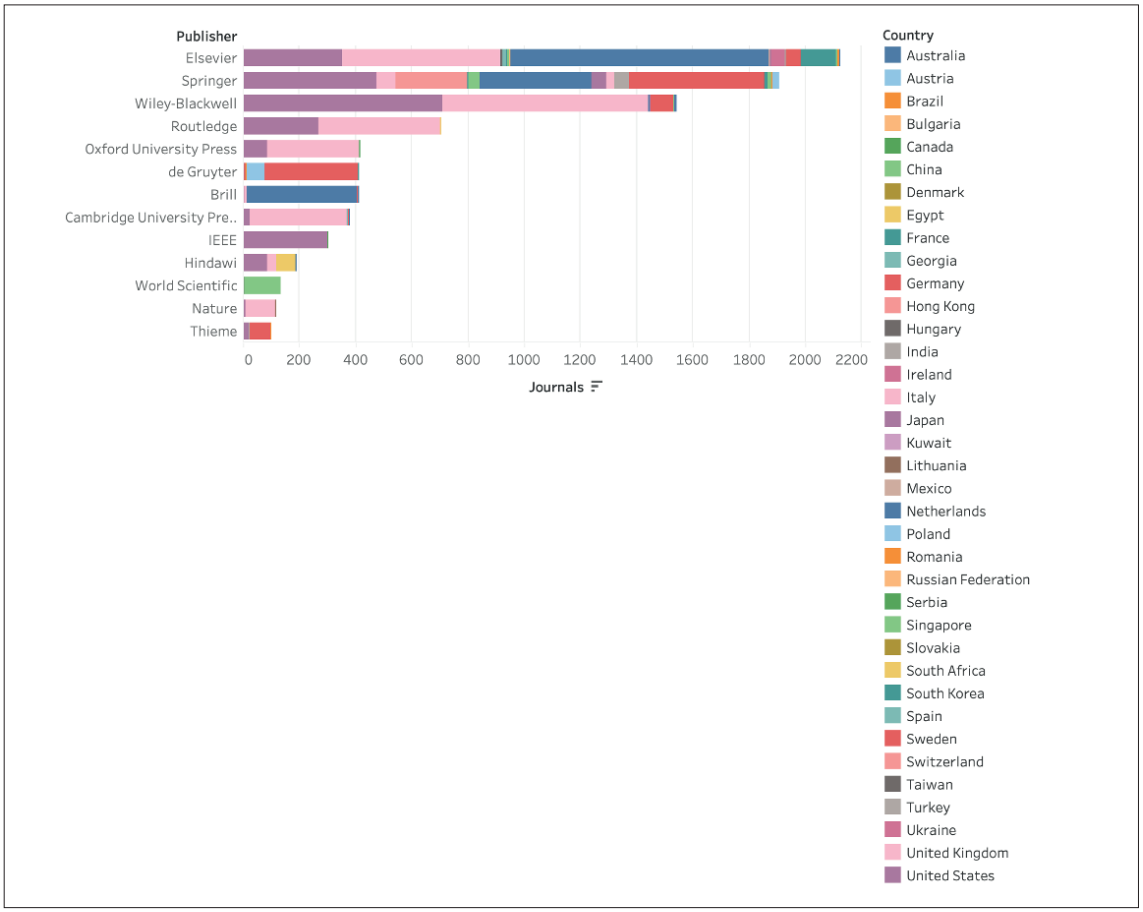

Figura 1. Volumen de títulos por países en cada editor.

5.1. Valoración de las respuestas

Cabe señalar que en ninguno de los 37 canales analizados aparece la marca de verificación que realiza Telegram de los canales oficiales, en muy pocos de ellos se añade una url, que es un elemento esencial en la verificación, observándose que la información publicada en los canales es poco homogénea. Ocurriendo algo parecido con el número de seguidores y el número de publicaciones de cada canal, que resultan muy dispares. En el análisis manual de los canales se concluyó que solo 8 de 37 eran canales reales y que estaban efectivamente vinculados con las editoriales, un 21,62 % del total. En la figura 2 vemos esos resultados.

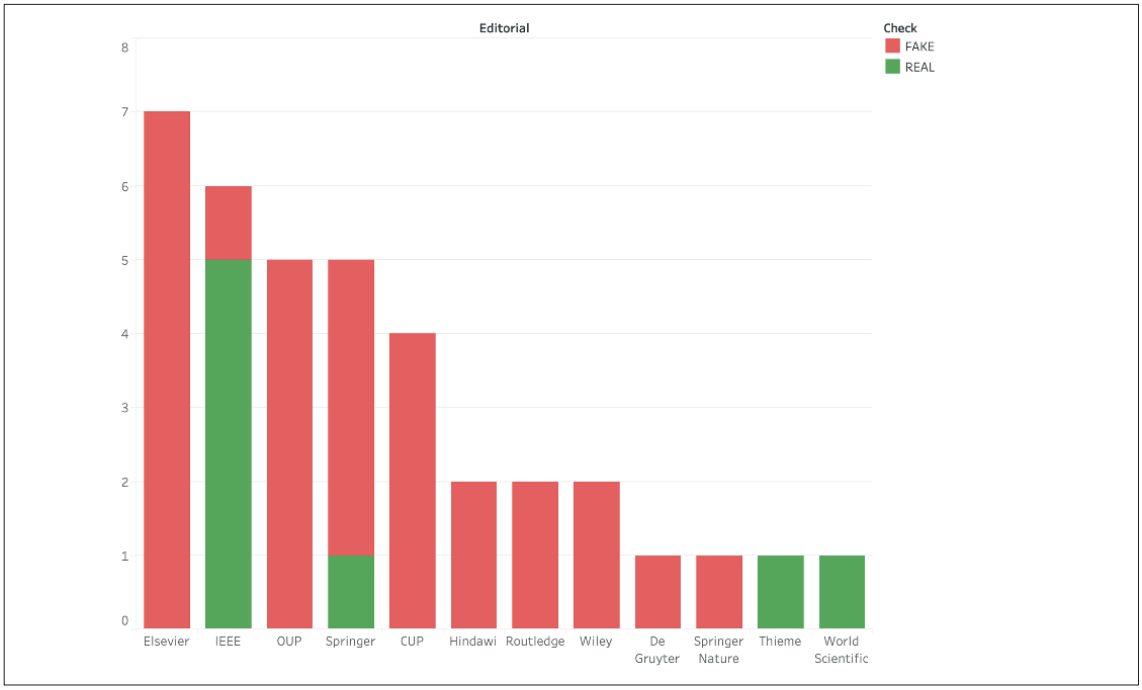

Figura 2. Relación de canales verdaderos y falsos por editorial

El análisis del contenido de los canales de Telegram, hecho por los modelos, revela conclusiones detalladas sobre lo que publican estos canales, lo cual es fundamental para determinar si son reales o falsos. En cuando al análisis y detección de contenidos fraudulentos, varios canales son considerados como FAKE basándose en el tipo de servicios y el lenguaje utilizado, que es inconsistente con las prácticas de las editoriales de prestigio, como anuncios repetidos para el envío de artículos, tarifas de publicación y plazos de envío rápidos. Llama la atención que aparezcan afirmaciones inusuales, como tiempos de revisión y correcciones de pruebas gratuitas e indexación en revistas por una tarifa, prácticas asociadas a revistas depredadoras.

Los modelos muestran acierto al detectar publicaciones con contenidos fraudulentos, y localizar afirmaciones contradictorias con las políticas de comercialización de las propias editoriales, apareciendo libros en diferentes formatos sin proporcionar información clara sobre derechos de autor o licencias. Los modelos también detectan presencias incongruentes, como la distribución de materiales de Cambridge University Press, a pesar de llevar el nombre de Oxford University Press.

En cuanto al contenido consistente con afiliación oficial o legítima, en algunos canales, principalmente aquellos afiliados a IEEE, localiza contenidos que respaldan (o al menos no refutan) su legitimidad local o temática. El análisis del contenido de los canales demuestra que la mayoría de los canales FAKE utilizan el contenido (libros gratuitos, promesas de publicación rápida) como señuelo para atraer usuarios, mientras que los canales clasificados como REALES o legítimamente afiliados (principalmente IEEE) se circunscriben a contenidos estrictamente profesionales, técnicos y vinculados a plataformas oficiales de la organización. Los modelos también tienen en cuenta, aunque en menor medida, el hecho de la inactividad de los canales o las escasas publicaciones de algunos.

En la revisión manual y en las respuestas de los modelos, en los canales FAKE se observa que hay una intencionalidad de suplantación, utilizando diferentes derivaciones en el nombre del canal y su identificación, con clara intención de suplantación de la editorial, en algunos casos para la difusión gratuita de publicaciones que se comercializan en webs oficiales. En cuanto a los usuarios y el número de publicaciones, son numerosos los canales que tienen escasas publicaciones y que se encuentran inactivos desde hace años.

También se observa que la mayoría de los correspondientes a IEEE son reales. Esto se debe al hecho de que IEEE es una asociación de profesionales además de ser una editorial, estos canales dedican sus publicaciones a cuestiones diferentes, y aunque hay referencias y enlaces a publicaciones, su creación y su uso no responde exclusivamente a finalidades editoriales.

Del mismo modo existe un reducido grupo de pequeñas editoriales, que utilizan el canal como un elemento de difusión de información sobre sus publicaciones y actividades diversas, pero el hecho de que no hayan gestionado con Telegram su check de veracidad, podría ser un indicativo de que el canal no es una prioridad en sus políticas de difusión y comercialización.

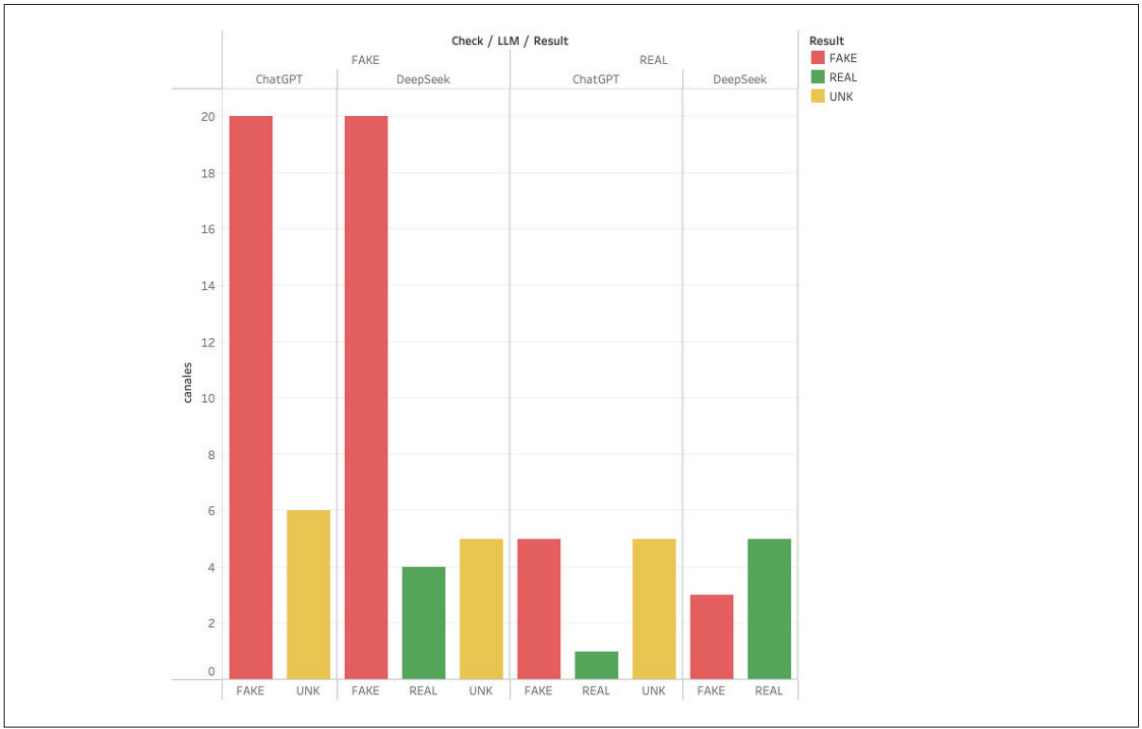

Figura 3. Relación de canales verdaderos y falsos por consideración de los modelos.

En la figura 3 se muestra un alto nivel de acierto de ambos modelos en la determinación de la falsedad de los canales FAKE. En el caso de ChatGPT no considera real ninguno de los falsos, considerando como dudosos un grupo semejante al considerado por DeepSeek como tales. Destaca el grupo de FAKE que DeepSeek considera reales, traslucién-dose que el envío desde los contenidos a dominios legítimos y de instituciones oficiales pueden haber sido causantes de la consideración como reales.

En el caso de la respuesta ante los canales reales se observa que DeepSeek reparte sus respuestas entre reales y falsos, no considerando dudoso a ninguno, mientras que ChatGPT limita el número de los que considera reales a uno, considerando dudosos o falsos a un grupo semejante. Esto muestra que los modelos tienden a fallar en la identificación de canales reales debido a la dependencia de criterios de verificación global (insignia azul y enlaces corporativos), ignorando la legitimidad del contenido o la afiliación regional.

5.2. Sitios referenciados

En el dataset podemos encontrar los rankings de los dominios más utilizados por cada modelo. Hay un total de 217 sitios web que fueron citados un total de 507 veces. ChatGPT ha hecho un uso más intenso de las referencias, un total 319 frente a 188 de Deepseek. Como la cantidad de sitios es muy grande, en la tabla X solo vemos a los más frecuentes de cada LLM.

ChatGPT				Deepseek			
Sitio	**Grupo**	**País**	**#**	**Sitio**	**Grupo**	**País**	**#**
t.me	Telegram	AE	35	t.me	Telegram	AE	33
facebook.com	Redes sociales	US	15	keepersecurity.com	Ciberseguridad	US	11
telemetr.io	Telegram	US	9	telegram.org	Telegram	AE	9
youtube.com	Redes sociales	US	9	infostealers.com	Ciberseguridad	US	7
reddit.com	Redes sociales	US	9	lifelock.norton.com	Ciberseguridad	US	6

Tabla 2 – Top 5 de sitios mencionados

Para poder analizar a todos los sitios los hemos agrupado en categorías y además le hemos incluidos el país de referencia. En algunos casos esto ha sido sencillo, pero en otros no tanto. Además de revisar los sitios web, hemos usado el servicio Whois e incluso le hemos preguntado a los propios LLMs. De esta manera obtenemos el ranking de la tabla 2, donde tenemos los cada grupo y la frecuencia acumulada en la que aparecen (ChatGPT + Deepseek). El grupo principal es el que constituye las diferentes variantes del servicio Telegram. Es razonable que así sea ya que en el propio prompt le hemos pedido a los modelos que analicen este servicio. En segundo lugar tenemos los servicios de redes sociales en general, algunos tan conocidos como Facebook, Youtube o Reddit, pero también hay otros menos destacados. Dentro de Editores top incluimos los sitios web de los editores objeto de estudio, y dentro de Editores a los restantes.

Grupo	**Frecuencia**
Telegram	117
Redes sociales	64
Editores top	52
Ciberseguridad	50
Repositorios	41
Editores	39
Universidades	37
Medios	24
Gobiernos	22
Marketing digital	17
Black hat académico	16
Crypto	15
Blog educativo	7
Revistas académicas	3
Fake news	2
Marketplace	1

Tabla 3 – Frecuencia de sitios por grupo

Es importante destacar el grupo Ciberseguridad ya que ha habido una gran cantidad de sitios relacionados con esta temática, a priori no esperable, y también ocurrió lo mismo con todo lo relacionado con el mundo Crypto. Del resto quizás es necesario destacar lo que hemos llamado Black hat académico y que no es más que sitios web donde se facilita el acceso a material con derechos de autor en una clara (o disimulada) vulneración de estos.

En la figura 4 podemos apreciar los sesgos de cada modelo con cada grupo. Claramente Telegram aparece en ambos, sin embargo, Ciberseguridad (y Cryto en menor medida) destaca en DeepSeek. ChatGPT, por su parte parece utilizar sitios más mainstreams como las propias redes sociales, las universidades o sitios del gobierno. Los Editores, Editores top y Repositorios aparecen referenciados por ambos modelos. Cabe destacar que el Black hat académico es mencionado por ambos modelos aunque ChatGPT lo hace en mayor medida.

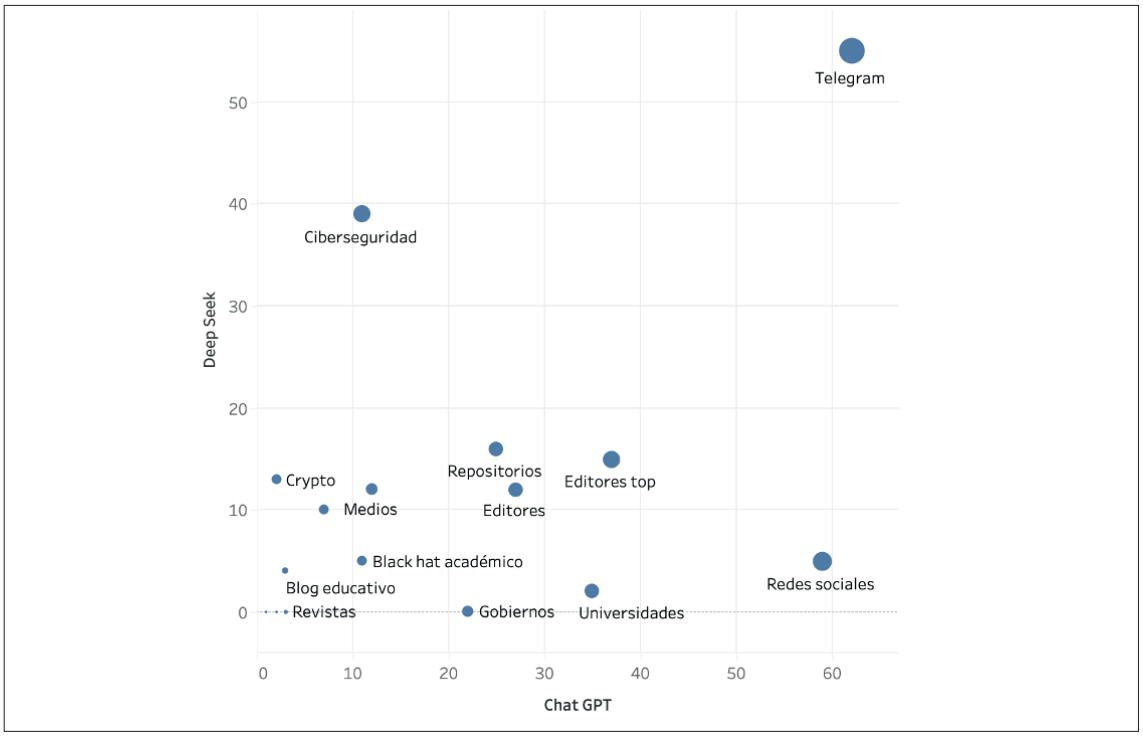

Figura 4. Comparación de sitios por modelo

Pero lo que más llama la atención es la procedencia de los medios por países. En este caso no se aprecia ningún sesgo destacable. En la figura 5 aparecen representados los países con dos más apariciones. El país predominante es Estados Unidos, tanto para el modelo estadounidense como para el chino. En segundo lugar, tenemos a Emiratos Arabes ya que es la sede del servicio Telegram, y un poco más atrás Reino Unido y Alemania. Hasta acá un dominio total del espacio web occidental. En quinto y sexto lugar tenemos a dos actores no occidentales: Rusia e India. Llama la atención que Rusia aparezca sesgado hacia ChatGPT mientras que la India lo hace hacia DeepSeek, donde especialmente lo ha usado dentro de la temática de la ciberseguridad.

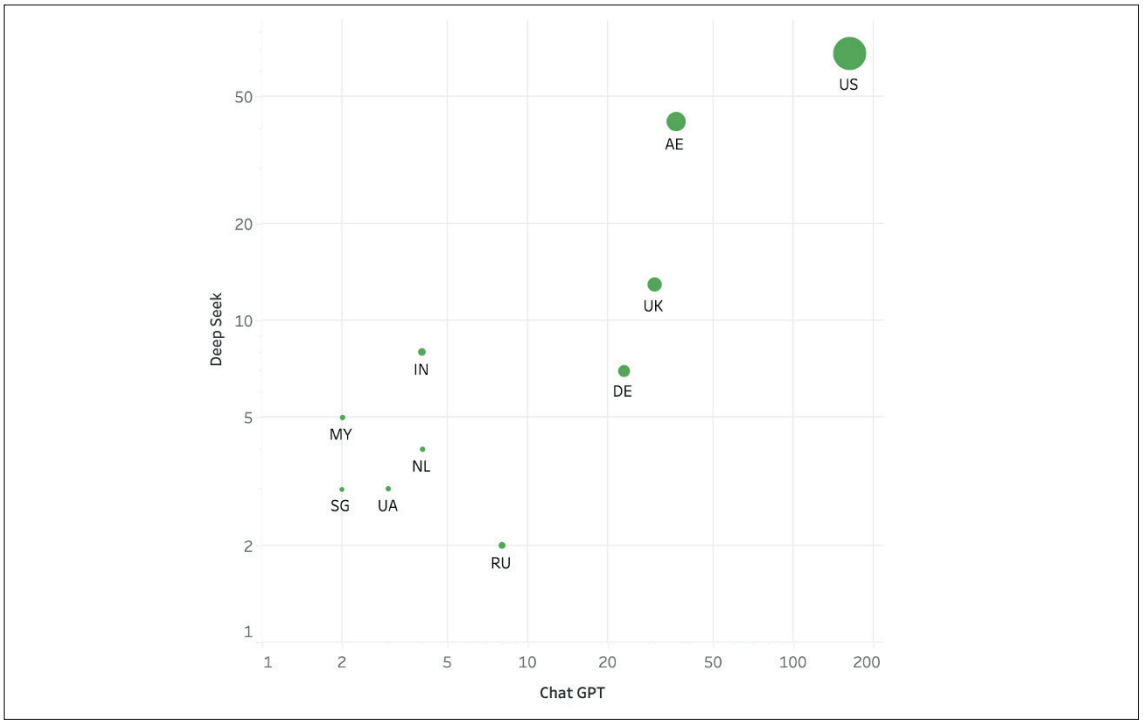

Figura 5. Distribución de países por modelo

Por último, debemos destacar que en todo el dataset solo hay un sitio chino, algo realmente curioso.

6. Discusión

La presente investigación confirma la proliferación masiva de canales no oficiales que suplantan la identidad de las principales editoriales académicas en Telegram. Los datos revelan que el 78,38 % de los canales analizados presentan características fraudulentas o prácticas no autorizadas, constituyendo un ecosistema paralelo que opera al margen de los canales oficiales de comunicación científica. Este hallazgo es consistente con investigaciones previas que han documentado la existencia de una extensa red de canales falsos y clones en la plataforma (La Morgia et al., 2021; 2023), y establece que la suplantación de identidad o la operación de canales no oficiales es la norma, no la excepción, en este ecosistema.

Este fenómeno no puede disociarse del contexto más amplio de la hegemonía editorial occidental identificada en nuestra Figura 1, donde Estados Unidos, Reino Unido y Países Bajos concentran la mayor parte de la producción. Esta concentración crea inevitablemente barreras de acceso económicas, idiomáticas y geopolíticas, que a su vez generan un caldo de cultivo para mercados informales que prometen eludir dichas restricciones. La aparición recurrente en nuestras fuentes de categorías como "black hat académico" corrobora la existencia de una infraestructura digital organizada que sustenta esta economía gris, un fenómeno que ha sido identificado también en el contexto del fraude editorial académico (Stockemer y Reidy, 2024).

En cuanto a los mecanismos de fraude identificados, los canales FAKE utilizan tácticas específicas inconsistentes con la comunicación académica legítima. En primer lugar, la promoción fraudulenta, caracterizada por un lenguaje informal y promocional, con afirmaciones inusuales como plazos de publicación extremadamente rápidos (2-3 semanas) y la mención de tarifas elevadas ($1200 USD), que se alinean con las características de revistas depredadoras o estafas. En segundo lugar, la distribución no autorizada y piratería: los canales relacionados con editoriales de libros (OUP, Springer) ofrecen acceso gratuito a materiales que el editor normalmente comercializa, lo cual resulta atípico para un editor legítimo. Estos patrones coinciden con los descritos por (Herasimenka et al. 2023) en su análisis de plataformas con moderación mínima, donde la facilidad de creación de canales y la ausencia de verificación preventiva facilitan la proliferación de contenidos engañosos. La característica más común de los canales no oficiales es la ausencia del sello de verificación azul de Telegram, que editoriales de la talla de las que estamos estudiando deberían ostentar si el canal fuera oficial.

La gran presencia de canales FAKE, que a menudo combinan la distribución de contenido pirata con servicios de publicación sospechosos, representa un riesgo significativo para la integridad académica y la marca de los editores. Este hallazgo viene a ratificar que la posibilidad de generar cualquier canal público, con cualquier denominación, al estar prevista la verificación solo *a posteriori* y únicamente si la corporación o entidad la solicita, abre un espacio notable a las posibilidades de fraude. Resulta llamativo que las editoriales científicas no parezcan mostrar un interés especial por esta vía que, bien explotada y correctamente verificada, podría constituir un medio de difusión robusto para las mismas. Esta desatención institucional contrasta con la creciente relevancia de Telegram como plataforma de difusión informativa (Willaert, 2023; Sánchez y Martos, 2020).

Respecto a la idoneidad de los modelos de LLM como herramientas de análisis, nuestros resultados presentan un panorama matizado. La significativa coincidencia entre ambos modelos en la identificación de canales claramente fraudulentos sugiere que estas he-

rramientas pueden efectivamente detectar patrones comunes de suplantación, lo cual es consistente con investigaciones recientes que demuestran la capacidad de los LLMs para identificar contenidos engañosos mediante análisis contextual (Chen y Shu, 2024; Papageorgiou et al., 2025). No obstante, las notables discrepancias en la verificación de los canales reales revelan limitaciones estructurales profundas. DeepSeek demuestra una tendencia marcada hacia lo que podríamos denominar un "criterio contextual", mostrando mayor disposición a validar canales que presentan contenido coherente y colaboraciones locales, incluso ante la ausencia de verificación formal. Por contraste, ChatGPT opera bajo un paradigma de "verificación estricta", requiriendo evidencias explícitas de afiliación institucional y mostrando una cautela considerablemente mayor.

Esta divergencia en los criterios de evaluación tiene implicaciones importantes. Los modelos pueden ser engañados por la apariencia profesional del contenido o por el uso de términos técnicos (como Scopus, Nature, etc.), sin poder validar los indicadores críticos de autenticidad. Esta limitación conecta con las advertencias sobre las alucinaciones de los LLMs —la generación de afirmaciones plausibles pero falsas— que comprometen su fiabilidad como detectores autónomos (Bender et al., 2021; Pooley, 2024). Los resultados refuerzan la necesidad de considerar estas herramientas como apoyo complementario al juicio experto, no como sustitutos del mismo, tal como sugieren (Patel et al., 2025) en su propuesta de sistemas de detección de desinformación.

El examen de las fuentes web consultadas por los modelos revela arquitecturas de conocimiento notablemente diferentes. Más allá del esperable predominio de dominios asociados a Telegram, emerge con fuerza inusitada el ámbito de la ciberseguridad en el caso de DeepSeek, mientras que ChatGPT muestra preferencia por fuentes institucionales y redes sociales convencionales. Esta disparidad sugiere que cada modelo construye su marco interpretativo sobre agentes de búsqueda web sustancialmente distintos, lo que necesariamente condiciona sus evaluaciones. DeepSeek, al apoyarse en recursos especializados en seguridad digital, parece desarrollar una sensibilidad aguda hacia patrones de fraude y suplantación, un enfoque que recuerda a las técnicas de detección basadas en análisis de patrones descritas por (Shu et al., 2017). ChatGPT, con su base más amplia y convencional, prioriza la verificación institucional formal. Ambos enfoques presentan ventajas complementarias, pero su desconexión explica en gran medida las discrepancias observadas en las clasificaciones.

El predominio absoluto de fuentes occidentales en ambos modelos refleja la hegemonía aún vigente de los aún llamados países en la arquitectura global de internet, un sesgo que ha sido documentado en estudios previos sobre la distribución geográfica del conocimiento en línea (Allcott y Gentzkow, 2017). Sin embargo, las afinidades diferenciales —DeepSeek con fuentes indias y sudasiáticas, ChatGPT con rusas— sugieren que los agentes de búsqueda de fuentes no están completamente predeterminados. El caso particularmente elocuente de la casi ausencia de fuentes chinas, a pesar del origen de DeepSeek, merece investigación adicional. Este hallazgo conecta con las observaciones de (Urman y Katz, 2022) sobre cómo las estructuras de red en Telegram pueden replicar divisiones ideológicas y geográficas más amplias.

Finalmente, nuestros resultados tienen implicaciones para la comprensión del fenómeno más amplio del desorden informativo descrito por (Wardle y Derakhshan, 2017). Los canales falsos de Telegram que suplantan a editoriales académicas no constituyen únicamente un problema de propiedad intelectual o fraude comercial, sino que erosionan la confianza en las instituciones de comunicación científica. En un contexto donde la desinformación científica ha demostrado tener consecuencias graves para la salud pública y el debate democrático (Díez-Garrido et al., 2021; Prieto-Campo et al., 2024), la proliferación de estos canales representa una amenaza estructural que requiere respuestas coordinadas tanto de las plataformas tecnológicas como de las propias editoriales académicas.

7. Conclusiones

Este trabajo pone de manifiesto que, en el ecosistema de Telegram, la presencia de las grandes editoriales científicas está profundamente distorsionada por una mayoría abrumadora de canales no oficiales que suplantan su identidad. A partir del análisis de 37 canales asociados a 13 editoriales líderes, se constató que el 78,38 % de los canales son fraudulentos o carecen de vínculos verificables con la editorial correspondiente, mientras que solo un 21,62 % puede considerarse real o legítimamente afiliado. Esta situación configura un entorno estructuralmente propenso a la desinformación y a la vulneración de derechos de propiedad intelectual.

En relación con la primera pregunta de investigación, nuestros resultados confirman la existencia de una red extensa y organizada de canales no oficiales que se apropian de nombres, logotipos y descripciones propias de las editoriales académicas. Estos canales combinan la distribución no autorizada de contenidos (especialmente libros) ofertas de servicios editoriales de dudosa legitimidad, alineados con prácticas propias de revistas depredadoras. La ausencia generalizada de verificación oficial de Telegram y la limitada presencia activa de las editoriales en la plataforma favorecen que estos canales fraudulentos se conviertan, de facto, en referentes visibles para muchos usuarios.

En cuanto a la segunda pregunta de investigación, los resultados muestran que ChatGPT y DeepSeek son herramientas útiles para el mapeo y la caracterización inicial de este ecosistema fraudulento, especialmente en la identificación de canales claramente FAKE. Ambos modelos coinciden con frecuencia en la detección de patrones de fraude. Sin embargo, presentan limitaciones importantes en la validación de canales reales, donde tienden a sobrerreaccionar ante la ausencia de señales institucionales fuertes (canales verificados, enlaces web corporativos) y a infravalorar contenidos legítimos con visibilidad local o regional. La diferencia entre el enfoque más "contextual" de DeepSeek y la mayor exigencia de "verificación formal" de ChatGPT sugiere que estos modelos operan bajo lógicas de evaluación parcialmente divergentes, lo que refuerza la necesidad de utilizar LLMs como apoyo complementario, y no como sustituto, del juicio experto humano.

Con relación a la tercera y cuarta pregunta, la temática de las fuentes consultadas revela que la construcción de conocimiento en estos sistemas está mediada por arquitecturas documentales específicas que reflejan prioridades temáticas y alineamientos diferenciados. El sesgo hacia la ciberseguridad de DeepSeek y el "institutionalismo" (universidades, gobiernos, editores) de ChatGPT no parecen constituir meras diferencias técnicas, sino visiones complementarias de un fenómeno complejo. Sin embargo, esta diferencia no se verifica claramente cuando miramos el origen de las fuentes. Existe un sesgo hacia fuentes occidentales, lo cual parece indicar que los entrenamientos de ambos modelos se han basado en conjuntos de datos no muy diferentes. La presencia de fuentes chinas es casi nula (0,2%), incluso en el caso de DeepSeek, un modelo que por momentos da la impresión de ser un indio o del sudeste asiático (Malasia y Singapur). En conjunto, los resultados sugieren que las diferencias entre ambos LLMs se explican más por la lógica de sus agentes de búsqueda y la priorización temática de sus fuentes que por el país de origen de cada modelo.

Como líneas de trabajo futuras, este estudio abre varias vías de interés. En el plano metodológico, resulta pertinente incorporar otros LLMs relevantes. (De ellos destacan actualmente, según nuestro criterio, Claude (Anthropic), Gemini (Google), Qwen (Alibaba), el controvertido Grok (X) y el recién llegado Kimi K2. Con ellos podremos evaluar si las diferencias observadas en este trabajo se sostienen en un espectro más amplio de modelos. En el plano aplicado, sería valioso extender el enfoque a otros tipos de desinformación científica y política presentes en Telegram, incluyendo la detección de *fake*

news y narrativas conspirativas. La progresiva integración de capacidades de análisis masivo de datos por parte de los LLMs ofrece una oportunidad para diseñar sistemas híbridos en los que la escala computacional y el criterio humano se refuercen mutuamente en la protección de la integridad de la comunicación científica.

Agradecimientos

Este paper ha recibido financiación del programa Proyectos de Generación del Conocimiento 2023 del Ministerio de Ciencia, Innovación y Universidades de España por el proyecto "Inteligencia Artificial en Europa ¿Auge o declive?" (PID2023-149646NB-I00).

8. Bibliografía

Abu-Ayfah, Zainab A. 2020. "Telegram App in learning English: EFL students' perceptions". *English Language Teaching*, 13(1): 51-62. https://doi.org/10.5539/elt.v13n1p51

Allcott, Hunt, & Matthew Gentzkow. 2017. "Social media and fake news in the 2016 election". *Journal of Economic Perspectives*, 31(2): 211-36. https://doi.org/10.1257/jep.31.2.211

Bender, Emily M., Timnit Gebru, Angelina McMillan-Major, & Shmargaret Shmitchell. 2021. "On the dangers of stochastic parrots: can language models be too big?". In *Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency*, 610-623. https://doi.org/10.1145/3442188.3445922

Carew, Sinéad, Amanda Cooper, & Ankur Banerjee. 2025. "DeepSeek sparks global AI selloff; Nvidia loses about $593 Billion of value". *Reuters*, January 27. https://www.reuters.com/technology/chinas-deepseek-sets-off-ai-market-rout-2025-01-27

Chen, Caiwei. 2025. "How a top Chinese AI model overcame US sanctions". *MIT Technology Review*, January 24. https://www.technologyreview.com/2025/01/24/1110526/china-deepseek-top-ai-despite-sanctions

Chen, Canyu, & Kai Shu. 2024. Combating misinformation in the age of LLMs: Opportunities and challenges. *AI Magazine*, 45(3): 313-331. https://doi.org/10.1002/aaai.12188

Cho, Hichang, Sungjong Roh, & Byungho Park. 2019. "Of promoting networking and protecting privacy: Effects of defaults and regulatory focus on social media users' preference settings". *Computers in Human Behavior*, 101. https://doi.org/10.1016/j.chb.2019.07.001

Cisternas-Osorio, Rodrigo, Alberto J. López-Navarrete, Margarita Cabrera-Méndez, & Rebeca Díez-Somavilla. 2022. "Telegram para el ejercicio de la comunicación interna: Análisis de su uso en universidades hispanohablantes". *Fonseca, Journal of Communication*, 25: 77-93. https://doi.org/10.14201/fjc.29750

Dargahi Nobari, Arash, Negar Reshadatmand, & Mahmood Neshati. 2017. "Analysis of Telegram, an instant messaging service". In *Proceedings of the 2017 ACM on Conference on Information and Knowledge Management*, 2035-2038. https://doi.org/10.1145/3132847.3133132

Dargahi Nobari, Arash, Malikeh Haj Khan Mirzaye Sarraf, Mahmood Neshati, & Farnaz Erfanian Daneshvar. 2021. "Characteristics of viral messages on Telegram; The world's largest hybrid public and private Messenger". *Expert Systems with Applications*, 168: 114303. https://doi.org/10.1016/j.eswa.2020.114303

Díez-Garrido, María, Covadonga Renedo Farpón, & Lorena Cano-Orón. 2021. "La desinformación en las redes de mensajería instantánea. Estudio de las fake news en los canales relacionados con la ultraderecha española en Telegram". *Miguel Hernández Communication Journal*, 12(2): 467-89. https://doi.org/10.21134/mhjournal.v12i.1292

Durov, Pavel. 2024. Canal Telegram Pavel Durov, 6 de septiembre. https://t.me/PavelDurovs/284

Eisenhardt, Kathleen M. 1989. "Building theories from case study research". *Academy of Management Review* 14(4): 532-50. https://doi.org/10.2307/258557

Flores-Vivar, Jesús Miguel, & Francisco José García-Peñalvo. 2023. "Reflexiones sobre la ética, potencialidades y retos de la Inteligencia Artificial en el marco de la Educación de Calidad (ODS4)". *Comunicar*, 31(74): 37-47. https://doi.org/10.3916/C74-2023-03

García-Marín, David. 2024. "Periodismo contra la desinformación. Proceso y estructura de las verificaciones en el fact-checking". *Infonomy*, 2(2): e24026. https://doi.org/10.3145/infonomy.24.026

Ghaffari, Mohtasham, Sakineh Rakhshanderou, Yadollah Mehrabi, & Afsson Tizvir. 2017. "Using social network of Telegram for education on continued breastfeeding and complementary feeding of children among mothers: A successful experience from Iran". *International Journal of Pediatrics*, 5(7): 5275-86. https://doi.org/10.22038/ijp.2017.22849.1915

Gregorio, Jesús, Alfredo Gardel, & Bernardo Alarcos. 2017. "Forensic analysis of Telegram messenger for Windows phone". *Digital Investigation*, 22: 88-106. https://doi.org/10.1016/j.diin.2017.07.004

Gregorio, Jesús. 2000. *Contribuciones al análisis forense de evidencias digitales procedentes de aplicaciones de mensajería instantánea*. Tesis doctoral, Universidad de Alcalá. https://ebuah.uah.es/dspace/handle/10017/50754?locale-attribute=es

Herasimenka, Aliaksandr, Jonathan Bright, Aleksi Knuutila, & Philip N. Howard. 2023. "Misinformation and professional news on largely unmoderated platforms: The case of Telegram". *Journal of Information Technology & Politics*, 20 (2): 198-212. https://doi.org/10.1080/19331681.2022.2076272

Herrero-Solana, Víctor, & Carlos Castro-Castro. 2022. "Telegram channels and bots: A ranking of media outlets based in Spain". *Societies*, 12(6): 164. https://doi.org/10.3390/soc12060164

Jakobsen, Jakob Bjerre. 2015. *A practical cryptanalysis of the Telegram messaging protocol*. Master's thesis, Aarhus University. https://enos.itcollege.ee/~edmund/materials/Telegram/A-practical-cryptanalysis-of-the-Telegram-messaging-protocol_master-thesis.pdf

Jalilvand, Asal, & Mahmood Neshati. 2020. "Channel retrieval: Finding relevant broadcasters on Telegram". *Social Network Analysis and Mining*, 10(1). https://doi.org/10.1007/s13278-020-0629-z

Kayaalp, Mahmut E., Robert Prill, Erdem A. Sezgin, Ting Cong, Aleksandra Królikowska, & Michael T. Hirschmann. 2025. "DeepSeek versus ChatGPT: Multimodal artificial intelligence revolutionizing scientific discovery. From language editing to autonomous content generation. Redefining innovation in research and practice". *Knee Surgery, Sports Traumatology, Arthroscopy*, 33(5), 1553-1556. https://doi.org/10.1002/ksa.12628

Kitsa, Mariana. 2023. "Telegram news channels: Overview of audience preferences and their implications". *Social Journal Studies*, 12(6): 354-69.

La Morgia, Massimo, Alessandro Mei, Alberto Maria Mongardini, & Jie Wu. 2018. "Pretending to be a VIP! Characterization and detection of fake and clone channels on Telegram". *ACM Transactions on the Web*, 12(4). https://dl.acm.org/doi/pdf/10.1145/3705014

La Morgia, Massimo, Alessandro Mei, Alberto Maria Mongardini, & Jie Wu. 2021. "Uncovering the Dark Side of Telegram: Fakes, clones, scams, and conspiracy movements". arXiv preprint arXiv:2111.13530. https://arxiv.org/pdf/2111.13530

La Morgia, Martina, Alessandro Mei, Alberto Maria Mongardini, & Jie Wu. 2023. "It's a trap! Detection and analysis of fake channels on Telegram". In *2023 IEEE International Conference on Web Services (ICWS)*, 97-104. IEEE. https://doi.org/10.1109/ICWS60048.2023.00026

Lewis, Patrick, Ethan Perez, Aleksandra Piktus, Fabio Petroni, Vladimir Karpukhin, Naman Goyal, Heinrich Küttler, et al. 2020. "Retrieval-augmented generation for knowledge-intensive NLP tasks". *Advances in Neural Information Processing Systems*, 33: 9459-9474.

Liu, Yinhan, Myle Ott, Naman Goyal, Jingfei Du, Mandar Joshi, Danqi Chen, Omer Levy, Mike Lewis, Luke Zettlemoyer, & Veselin Stoyanov. 2019. "RoBERTa: A Robustly Optimized BERT Pretraining Approach". *arXiv.* Preprint submitted July 26. https://arxiv.org/abs/1907.11692

López-Tárraga, Ana Belén. 2020."Comunicación de crisis y ayuntamientos: El papel de Telegram durante la crisis sanitaria de la Covid-19". *RAEIC: Revista de la Asociación Española de Investigación de la Comunicación*, 7(14): 104-26. https://doi.org/10.24137/raeic.7.14.5

Lu, Ying & Naiwei Yao. 2025. "A fake news detection model using the integration of residual networks and attention mechanisms". *Scientific Reports*, 15: 20544. https://doi.org/10.1038/s41598-025-05702-w

Martos Moreno, Javier, & Hada M. Sánchez Gonzales. 2024. "Consumo incidental de noticias en Telegram". *Estudios sobre el Mensaje Periodístico*, 30(1): 167-176. https://doi.org/10.5209/esmp.92127

Mohammed, Ibrahim A., Ibrahim I. Kuta, Oluwole C. Falode, & Ahmed Bello. 2024. "Comparative performance of undergraduate students in micro teaching using Telegram and WhatsApp in collaborative learning settings". *Journal of Mathematics and Science Teacher*, 4(2). https://doi.org/10.29333/mathsciteacher/14411

Ouyang, Long, Jeff Wu, Xu Jiang, Diogo Almeida, Carroll L. Wainwright, Pamela Mishkin, Chong Zhang, et al. 2022."Training language models to follow Instructions with human feedback". *Advances in Neural Information Processing Systems*, 35: 27730–44.

Papageorgiou, Eleftheria, Iraklis Varlamis, & Christos Chronis. "Harnessing Large Language Models and deep neural networks for fake news detection". *Information*, 16(4): 297. https://doi.org/10.3390/info16040297

Patel, Jainee, Chintan M. Bhatt, Himani Trivedi, Thanh Thi Nguyen, Kadi Sarva Vishwavidyalaya. 2025. "Misinformation detection using large Language models with explainability." In *Proceedings of the 8th International Conference on Algorithms, Computing and Artificial Intelligence*. https://arxiv.org/pdf/2510.18918

Prieto-Campo, Álvaro, Olalla Vázquez-Cancela, Fátima Roque, María-Teresa Herdeiro, Adolfo Figueiras, & Maruxa Zapata-Cachafeiro. 2024. "Unmasking vaccine hesitancy and refusal: A deep dive into anti-vaxxer perspectives on Covid-19 in Spain". *BMC Public Health*, 24(1751). https://doi.org/10.1186/s12889-024-18864-5

Pooley, Jeff. 2024. "Publicación en Large Language Model (LLM)". *SciELO in Perspective* (blog), 19 de enero. https://blog.scielo.org/es/2024/01/19/publicacion-en-llm

Sánchez Gonzales, Hada M., & Juan Martos Moreno. 2020. "Telegram como herramienta para periodistas: Percepción y uso". *Revista de Comunicación* 19(2): 245-61. https://doi.org/10.26441/RC19.2-2020-A14

Sedano Amundarain, Jon, & María Bella Palomo Torres. 2018. "Aproximación metodológica al impacto de WhatsApp y Telegram en las redacciones". *Hipertext.net*, 16. https://doi.org/10.31009/hipertext.net.2018.i16.10

Shu, Kai, Amy Sliva, Suhang Wang, Jiliang Tang, & Huan Liu. 2017. "Fake news detection on social media: A data mining perspective". *ACM SIGKDD Explorations Newsletter*, 19(1): 22-36.

StatCounter. 2025. "Leading AI chatbots on desktop devices worldwide between March and October 2025, by share of web referrals". Chart, 9 de noviembre de 2025. En *Statista*. https://www.statista.com/statistics/1625095/desktop-ai-chatbots-market-share

Steblyna, Nataliia O. 2024. "Digital media environment in wartime. Russian invasion coverage in Ukrainian professional and amateur news media". *Horyzonty Polityki*, 15(51): 99-119. https://doi.org/10.35765/hp.2484

Stockemer, Daniel, & Theresa Reidy. 2024. "Academic misconduct, fake authorship letters, cyber fraud: Evidence from the International Political Science Review". *Learned Publishing*, 37(1): 39-43. https://doi.org/10.1002/leap.1587

Sušánka, Tomáš, & Jozef Kokeš. 2017. "Security analysis of the Telegram IM". In *Proceedings of the 1st Reversing and Offensive-Oriented Trends Symposium*, 1-8. https://doi.org/10.1145/3150376.3150382

Telegram. "Mensajes por estrellas, regalos fijados, plataforma de verificación 2.0 y más". *Telegram*,12. https://telegram.org/blog/star-messages-gateway-2-0-and-more/es

Thorp, H. Holden. 2023. "ChatGPT is fun, but not an author". *Science*, 379(6630): 313. https://doi.org/10.1126/science.adg7879

Urman, Aleksandra, & Stefan Katz. 2022. "What they do in the shadows: Examining the far-right networks on Telegram". *Information, Communication & Society*, 25(7): 904-23. https://doi.org/10.1080/1369118X.2020.1803946

Vaswani, Ashish, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N. Gomez, Łukasz Kaiser, & Illia Polosukhin. 2017. "Attention is all you need". *Advances in Neural Information Processing Systems*, 30.

Wardle, Claire, & Hossein Derakhshan. 2017. *Information Disorder: Toward an Interdisciplinary Framework for Research and Policymaking*. Strasbourg: Council of Europe.

Willaert, Tom. 2023. A computational analysis of Telegram's narrative affordances. *Plos one*, 18(11), e0293508. https://doi.org/10.1371/journal.pone.0293508

Yin, Robert K. 2018. *Case study research and applications: Design and methods*. 6th ed. Los Angeles: SAGE Publications.

9. Apéndice I

Editoriales con sus variantes de denominación

Editorial	Variantes
Elsevier	Elsevier; Elsevier B.V. ; Elsevier (Singapore) Pte Ltd; Elsevier Editora Ltda; Elsevier Espana; Elsevier Espana S.L.U; Elsevier GmbH; Elsevier Inc.; Elsevier Ireland Ltd; Elsevier Ltd; Elsevier Masson s.r.l.; Elsevier Science; Elsevier Science & Technology; Elsevier Science B.V.; Elsevier Taiwan LLC; Elsevier USA; Reed Elsevier India Pvt. Ltd.; Reed-Elsevier (India) Private Limited
Springer	Springer; Springer Berlin; Springer Boston; Springer China; Springer Fachmedien Wiesbaden GmbH; Springer Gabler; Springer GmbH & Co, Auslieferungs-Gesellschaf; Springer Healthcare; Springer Heidelberg; Springer India; Springer International Publishing; Springer International Publishing AG; Springer Japan; Springer London; Springer Medizin; Springer Nature; Springer Netherlands; Springer New York; Springer Paris; Springer Publishing Company; Springer Science + Business Media; Springer Science and Business Media B.V.; Springer Science and Business Media Deutschland GmbH; Springer Science and Business Media, LLC; Springer Singapore; Springer US; Springer Verlag; Springer Vieweg; Springer-Verlag GmbH and Co. KG; Springer-Verlag Italia s.r.l.; Springer-Verlag Wien; SpringerOpen;
Wiley-Blackwell	John Wiley & Sons Inc; John Wiley and Sons Inc; John Wiley and Sons Ltd; Wiley-Blackwell; Wiley-Blackwell for the International Association for the Study of Obesity; Wiley-Blackwell Publishing Ltd; Wiley-Liss Inc.; Wiley-VCH Verlag;
Routledge	Routledge; Brunner - Routledge (US); Routledge, Taylor & Francis Group
Oxford University Press	Oxford University Press
De Gruyter	de Gruyter; De Gruyter Mouton; De Gruyter Oldenbourg; De Gruyter Open Ltd.; De Gruyter Saur; Walter de Gruyter; Walter de Gruyter GmbH; Walter de Gruyter GmbH & Co. KG
Brill	Brill Academic Publishers; Brill Nijhoff; Brill Schoningh; Brill: Rodopi
Cambridge University Press	Cambridge University Press
IEEE	IEEE Advancing Technology for Humanity; IEEE Canada; IEEE Circuits and Systems Society; IEEE Communications Society; IEEE Computational Intelligence Society; IEEE Computer Society; IEEE CONTROL SYSTEMS SOCIETY; IEEE Education Society; IEEE Electromagnetic Compatibility Society; IEEE Electron Devices Society; IEEE Industrial Electronics Society; IEEE Power Electronics Society; IEEE Systems, Man, and Cybernetics Society; Institute of Electrical and Electronics Engineers Inc.
Hindawi	Hindawi; Hindawi Limited; Hindawi Publishing Corporation; Wiley-Hindawi
World Scientific	World Scientific; World Scientific Publishing Co. Pte Ltd; World Scientific Publishing Co., Inc.
Nature	Nature Partner Journals; Nature Publishing Group; Nature Research; Nature Research Centre
Thieme	Thieme; Georg Thieme Verlag; Thieme Medical Publishers; Thieme Medical Publishers, Inc.; Thieme Publishers Rio

10. Apéndice II

Conclusiones de modelo por canal y revisión manual

Canal	Editorial	DeepSeek	ChatGPT	Check
@ScopusElsevier	Elsevier	FAKE	FAKE	FAKE
@scopusservices	Elsevier	FAKE	FAKE	FAKE
@ElsevierCentralAsia	Elsevier	REAL	UNK	FAKE
@Elsevierscienceuzbekistan	Elsevier	FAKE	FAKE	FAKE
@clinicalkey	Elsevier	FAKE	UNK	FAKE
@elsevier_iran	Elsevier	REAL	UNK	FAKE
@elsevieducatiRealn	Elsevier	UNK	UNK	FAKE
@springer_uzb	Springer	FAKE	FAKE	FAKE
@NatureClimateTelegram	Springer	FAKE	UNK	REAL
@MasterTez	Springer	FAKE	FAKE	FAKE
@TAQUspringer	Springer	FAKE	FAKE	FAKE
@SPBooksMed	Springer	FAKE	FAKE	FAKE
@wileyrus	Wiley	REAL	FAKE	FAKE
@wiley_uz	Wiley	UNK	FAKE	FAKE
@routledgebooks	Routledge	UNK	UNK	FAKE
@routledge_bobil	Routledge	FAKE	FAKE	FAKE
@oupbooks	OUP	FAKE	FAKE	FAKE
@oxford_ielts_grammar_booksN1	OUP	FAKE	FAKE	FAKE
@oxford_university_press_bot	OUP	FAKE	FAKE	FAKE
@studycollegeTt	OUP	FAKE	FAKE	FAKE
@edupressUzbekistan	OUP	FAKE	UNK	FAKE
@degruyter	De Gruyter	UNK	UNK	FAKE
@cambridgeuni	CUP	FAKE	FAKE	FAKE
@cambridge_university_press_ielts	CUP	UNK	FAKE	FAKE
@cambridgeUniversityPresss	CUP	FAKE	FAKE	FAKE
@Cambridge_practice_discussion	CUP	FAKE	FAKE	FAKE
@ieeear	IEEE	FAKE	UNK	REAL
@aetel	IEEE/UPM	FAKE	FAKE	REAL
@IEEEbot	IEEE/UGR	FAKE	FAKE	FAKE
@ieeesiberia	IEEE	REAL	UNK	REAL
@IEEEIranSection	IEEE	REAL	REAL	REAL
@IEEESmartGrid	IEEE	REAL	UNK	REAL
@hindawi_books	Hindawi	REAL	FAKE	FAKE
@HindawiFoundationBooks	Hindawi	FAKE	FAKE	FAKE
@wspcsg	World Scientific	REAL	FAKE	REAL
@NaturePublishingGroup	Springer Nature	FAKE	FAKE	FAKE
@thiemepublishers	Thieme	REAL	FAKE	REAL

FAKE = falso / UNK = dudoso / REAL = real