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

ARTÍCLE ORIGINAL

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

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

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

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

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

Rebut: 10-09-2025
Acceptat: 07-11-2025

Citació recomanada

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

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.

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.

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.

1. Introducció

L'aplicació de missatgeria instantània Telegram ha superat fa temps la seva funció inicial de missatgeria entre persones, convertint-se en una tecnologia sòlida per a l'intercanvi global d'informació. A diferència de les xarxes socials clàssiques, els usuaris solen associar-la amb comunicació íntima (Dargahi et al., 2017), passant desapercebut que els canals d'aquest servei poden constituir una font substancial d'informació. La manca de control centralitzat de publicació de canals fa fàcil que l'usuari pugui veure's confós amb l'origen i responsabilitat d'aquests canals, tal com ja s'ha demostrat per al cas de la premsa espanyola (Herrero i Castro, 2022).

La proliferació de notícies falses i desinformació s'ha convertit en una preocupació significativa a l'era digital. Amb l'auge de les plataformes de xarxes socials i la creixent facilitat de difusió d'informació, la propagació de contingut enganyós planteja serioses amenaces a l'opinió pública, els processos democràtics i l'harmonia social (Shu et al., 2017). En aquest context, els models grans del llenguatge (LLM) han emergit com a eines prometedores per combatre la desinformació, gràcies al seu profund coneixement del món i les seves potents capacitats de raonament. Investigacions recents han demostrat que models preentrenats ajustats, tant models antics com BERT i RoBERTa com els LLM més avançats, poden assolir resultats notables en la detecció de notícies falses sense necessitat de dades auxiliars extenses (Flores-Vivar i García-Peñalvo, 2023). Aquests models han mostrat particular eficàcia mitjançant estratègies de prompting, especialment el raonament en cadena de pensament (chain-of-thought), que millora significativament el rendiment en tasques de verificació de fets (Chen i Shu, 2024).

Per al present estudi, es van seleccionar dos LLM representatius del panorama actual: ChatGPT i DeepSeek. L'elecció de ChatGPT es deu al fet de ser el primer a irrompre (i cridar molt l'atenció) a finals de 2022 (Thorp, 2023) i que actualment té encara una posició dominant al mercat de models de llenguatge. Segons diverses anàlisis de mercat, ChatGPT manté una quota superior al 80% entre els chatbots d'IA generativa, consolidant-se com l'LLM més utilitzat tant per usuaris individuals com per empreses (StatCounter, 2025). La seva àmplia adopció i reconeixement el converteixen en el referent natural per avaluar capacitats de detecció de contingut fraudulent.

Per la seva part, la inclusió de DeepSeek obeeix a la seva singular rellevància en el moment de començar el disseny d'aquest experiment. El gener de 2025, aquesta startup xinesa va irrompre als mitjans internacionals en presentar el seu model R1, que va demostrar capacitats comparables a les dels seus competidors occidentals, però desenvolupat sota restriccions significatives d'accés a maquinari avançat. Les polítiques de control d'exportacions dels Estats Units, implementades des de l'octubre de 2022, van restringir severament a fabricants com Nvidia la venda dels seus xips més potents (H100 i A100) a empreses xineses. DeepSeek va aconseguir entrenar els seus models utilitzant xips H800, una versió amb rendiment limitat dissenyada específicament per al mercat xinès. En lloc de limitar-se a l'ús estàndard del framework de Nvidia (CUDA), els seus enginyers van recórrer a programació de molt baix nivell en PTX —el llenguatge tipus assemblador sobre el qual es recolza CUDA— per esprémer al màxim el potencial d'aquests xips (Chen, 2025). L'impacte de la seva aparició va generar fins i tot una forta caiguda borsària de la pròpia Nvidia: gairebé 600 mil milions de dòlars en un dia (Carew et al., 2025). Aquest cas representa un interessant contrapunt metodològic enfront de ChatGPT, en tractar-se d'un model desenvolupat sota restriccions tecnològiques que, no obstant això, assoleix nivells de rendiment competitius.

2. Marc teòric

2.1. Els canals de Telegram

L'èxit de WhatsApp com a alternativa als SMS quan es va començar a generalitzar l'ús de telèfons mòbils amb pantalla tàctil connectats a Internet va marcar l'inici del creixement de les aplicacions de missatgeria instantània. L'aparició de WeChat a la Xina i aplicacions com Telegram i Signal, presentades com a eines més transparents i segures, va fer néixer un entorn comunicatiu nou. Els diferents avenços de les diferents aplicacions han anat enriquint mútuament les capacitats de cadascuna d'elles, cosa que ha contribuït a un augment constant d'usuaris (Gregorio et al., 2017).

La incorporació dels canals a Telegram el 2015 va suposar el sorgiment d'una nova forma de comunicació unidireccional que permet difondre notícies, contingut d'entreteniment, comunicacions oficials, etc., a audiències globals il·limitades (Kitsa, 2023). La incorporació de novetats des d'aleshores ha estat una constant, com l'edició de missatges (que permet corregir errors o actualitzar informació després de la publicació), eines bàsiques d'anàlisi per rastrejar el rendiment dels missatges, millora de les capacitats de gestió i diversificació de formats i eines. Telegram no ha estat una aplicació estàndard, ja que algunes de les seves decisions sobre criptografia i seguretat són singulars (Jakobsen, 2015). La seva arquitectura de seguretat es presenta com un valor primordial per a la solidesa de la pròpia eina (Wardle i Derakhshan, 2017).

Telegram ha estat incorporat a l'activitat informativa en mitjans de comunicació (Sánchez i Martos, 2020), i s'ha comprovat la solidesa dels canals per a la difusió informativa (Sedano i Palomo, 2018). Telegram s'ha fet un lloc en el panorama mediàtic actual en establir-se com un canal alternatiu per al consum de notícies. S'ha demostrat que la confiança en la marca periodística juga un paper crucial en la selecció de les fonts d'informació en aquesta plataforma (Martos i Sánchez, 2024). La seva flexibilitat ha permès desenvolupar solucions informatives i comunicatives en situacions tan extraordinàries com la guerra d'Ucraïna, on s'han arribat a crear nous formats comunicatius (Steblyna, 2024). Utilitzats també per institucions universitàries d'Espanya i Llatinoamèrica per al seu ús en difusió de la seva informació (Cisternas et al., 2022). No ha ocorregut en la mateixa mesura en entorns purament comercials. Telegram s'ha mostrat sòlid per suportar sistemes d'informació documental i, al mateix temps, generar entorns comunicatius a diferents nivells. Aquestes capacitats s'han comprovat en diversos formats (Mohammed et al., 2024). La varietat dels mateixos, la solidesa del seu sistema d'emmagatzematge i la capacitat de difusió de continguts s'han usat amb èxit en experiències d'aprenentatge d'idiomes i reforçament de capacitats lingüístiques (Abu-Ayfah, 2020).

Telegram proporciona un model estructurat per al seu ús en la interacció social i la moderació experta (Ghaffari et al. 2017). Aquestes fortaleses no han donat els resultats desitjats en camps com el sanitari. El 2020, amb motiu de la pandèmia, des de Telegram es va realitzar un oferiment perquè les autoritats sanitàries de tot el món verifiquessin els seus canals oficials. Sorprenentment, la iniciativa només va ser seguida per una vintena de països i no va tenir un gran recorregut. Paradoxalment, en aquelles circumstàncies va ser l'eina escollida per molts organismes per difondre la seva informació en aquesta crisi (López, 2020). En aquella mateixa època, davant la demanda del teletreball, Telegram va millorar notablement les eines de videocomunicació, tant en xats com en grups i canals, assolint capacitats fins i tot majors que les d'algunes de les més populars plataformes de streaming, desenvolupant-se com un sistema híbrid que ha ampliat notablement les capacitats comunicatives (Dargahi et al., 2021).

Els canals de Telegram progressivament s'han convertit en veritables plataformes de difusió (Willaert, 2023). El perfeccionament dels permisos avançats per a administradors els dota de capacitats per assignar rols i gestionar de forma granular qui pot publicar, editar, fixar o eliminar missatges, convertint-los funcionalment en editors (Gregorio, 2020). La possibilitat afegida d'associar grups de discussió als canals, unida a la millora de les eines analítiques i estadístiques, ha permès a més disposar d'instruments per avaluar l'abast i la interacció de les publicacions (La Morgia et al., 2023).

El sistema de cerca de Telegram barreja la localització de canals per les paraules contingudes en el seu títol amb la localització de paraules en els continguts de tots els canals subscrits. Resulta molt ràpid i eficaç per a la localització de continguts en els canals subscrits, però molt confús i insegur en la localització precisa de canals (Jalilvand i Neshati, 2020).

En les actualitzacions de Telegram durant 2024 s'ha reforçat de manera notable tot el relatiu a les mini apps a la seva plataforma de bots, amb una creixent activitat al voltant dels premis i la monetització de serveis i activitats comercials. L'èxit d'aquests nous camps d'activitat ha obert una línia d'actuació que, en les primeres actualitzacions de 2025, s'està projectant sobre els canals. El llançament del sistema de verificació de tercers presentat per Telegram en la primera actualització de l'any és una bona prova d'això. Aquest nou sistema per combatre estafes i desinformació uneix al tradicional check la possibilitat que serveis oficials aliens al propi Telegram verifiquin comptes i xats. Els canals o comptes verificats mitjançant aquest mètode podran mostrar una icona única al costat del seu nom, i en fer clic, es podrà veure quin servei realitza la certificació i el motiu. En l'actualització de març de 2025, s'han afegit utilitats per al control de la safata d'entrada. Aquestes permeten establir una tarifa per als missatges entrants d'usuaris que no estiguin entre els contactes, cosa que facilita un control total sobre la safata d'entrada. Els usuaris podran guanyar estrelles (el sistema de pagaments utilitzat a les mini apps), filtrar els missatges no desitjats i evitar la sobrecàrrega de la safata d'entrada. Aquesta configuració també es pot aplicar als xats grupals i converses de canals per mantenir les interaccions enfocades i lliures d'spam, cosa que ajuda els seus propietaris a monetitzar la seva comunitat i guanyar estrelles a partir de converses significatives (Telegram, 2025). Totes aquestes mesures permeten augmentar la transparència i la confiança en el contingut publicat, però només donaran fruits si els mateixos mitjans, les institucions i la comunitat científica en fan ús.

Resulta indubtable que aquestes utilitats han obert oportunitats i noves vulnerabilitats, plantejant reptes als quals és necessari donar resposta (Sušánka i Kokeš, 2017). El fet que qualsevol pugui crear un canal públic de Telegram, imitar la imatge d'un original, anomenar-lo i fins i tot alimentar-lo amb missatges provinents de qualsevol font fa que el sistema no hagi mostrat solidesa, tota vegada que la verificació de canals no s'ha generalitzat. Encara que existeixen multitud de canals «teòricament oficials» fins i tot enllaçats des de les seves webs, molts apareixen a Telegram sense el check de verificació oficial (Herrero i Castro, 2022). De vegades apareixen canals amb denominacions iguals, en els quals resulta complicat verificar la seva naturalesa oficial, ja que fins i tot reprodueixen continguts provinents dels RSS dels mateixos mitjans i organitzacions (La Morgia et al., 2023). Hi ha evidència que les capacitats descrites permeten la fàcil difusió de desinformació a causa de la privacitat i la preservació de l'anonimat. Això ha donat com a resultat la proliferació de canals falsos, com es va evidenciar de manera notable després de la pandèmia (Díez et al., 2021).

En els últims anys, després de diversos episodis de ressò mediàtic, sembla que les relacions amb els organismes governamentals de diferents països estan sent més fluïdes i s'estan assolint més acords amb reguladors. Bona prova d'això són les eliminacions de milions de publicacions i canals nocius cada dia, la publicació d'informes diaris de transparència i les línies directes amb ONG per processar les sol·licituds de moderació amb més rapidesa (Durov, 2024).

La pròpia naturalesa dels canals, la seva facilitat de creació i la inexistent moderació (excepte la verificació o l'eliminació responent a una denúncia) fan que la difusió de desinformació sigui un fet inevitable (Herasimenka et al., 2023). El seu impacte és especialment preocupant en l'àmbit científic pels problemes de suplantació d'identitat mitjançant l'ús indegut de noms d'editorials i logos oficials per guanyar credibilitat instantània (Wardle i Derakhshan, 2017). La manipulació informativa mitjançant la difusió d'estudis falsos, manipulats o incomplets pot influir negativament en debats acadèmics i decisions polítiques (Cho et al., 2019). La inevitable erosió de la confiança, en confondre els usuaris, pot danyar la credibilitat de la ciència i les institucions acadèmiques (Allcott i Gentzkow, 2017).

La naturalesa dels canals de Telegram permet generar bases d'informació per analitzar-les i obtenir resultats fiables que poden ser avaluats. No obstant això, la tasca d'avaluació dels resultats de les anàlisis quantitatives ha requerit fins ara el concurs d'analistes humans (La Morgia et al., 2018); els LLM poden obrir una nova perspectiva en cas de poder prescindir dels humans per a aquesta tasca.

2.2. Irrupció dels LLM: ChatGPT i DeepSeek

El llançament de ChatGPT per OpenAI el novembre de 2022 va provocar una convulsió que va transcendir l'àmbit de les TIC, en posar la intel·ligència artificial generativa a l'abast de persones sense coneixements tècnics especialitzats. Milions d'usuaris van començar a interactuar amb la IA de manera senzilla, desencadenant una onada de canvis i debats que van afectar la societat a nivell global. Empreses tecnològiques líders a Silicon Valley van reaccionar ràpidament: Google va llançar Gemini, el seu model de llenguatge avançat; Microsoft va reforçar la seva aposta amb inversions milionàries en OpenAI; Meta va desenvolupar els seus propis models de llenguatge, com LLaMA, i el va oferir en obert; X (anteriorment Twitter) va presentar Grok, un model enfocat a la interacció social, entre molts altres. A més, empreses com Amazon i Apple també van intensificar els seus esforços en IA, integrant tecnologies generatives en els seus ecosistemes de serveis i dispositius. No obstant això, aquest auge es va veure abruptament interromput a principis de 2025 amb la irrupció de DeepSeek, una startup xinesa recolzada per un fons de cobertura. DeepSeek, una alternativa sorprenentment econòmica i de codi obert, no només va desafiar el domini d'OpenAI i altres empreses occidentals. DeepSeek va alterar gairebé totes les expectatives en el sector de la IA, demostrant que la innovació no està limitada a Silicon Valley i que el mercat global d'intel·ligència artificial és més dinàmic i competitiu del que molts anticipaven.

Més enllà de les implicacions socioeconòmiques, ChatGPT i DeepSeek representen dos enfocaments complementaris dins dels LLM, tecnologies que han redefinit el processament del llenguatge natural mitjançant l'arquitectura Transformer (Vaswani et al. 2017). Ambdós sistemes operen mitjançant sofisticats mecanismes d'atenció, que assignen rellevància contextual a cada terme, cosa que permet la identificació de patrons i relacions semàntiques en grans volums de dades, que, combinada amb sistemes de cerca a Internet, obre expectatives d'anàlisi d'informació molt prometedores.

ChatGPT s'ha consolidat com un referent en la generació de diàlegs fluids i contextualitzats. La seva eficàcia es basa en un procés en dues fases: un preentrenament massiu en corpus multilingües (llibres, articles, webs) i aprenentatge per reforç amb retroalimentació humana (RLHF) (Ouyang et al. 2022), on ajustadors humans qualifiquen respostes per prioritzar coherència i seguretat. Aquesta dualitat li permet adaptar-se a tasques tan diverses com la redacció creativa o la simulació d'assistència tècnica, amb un risc inherent: la generació d'al·lucinacions (afirmacions plausibles però falses) en extrapolar patrons de dades no verificades (Bender et al., 2021).

DeepSeek és un LLM de codi obert optimitzat per a l'anàlisi semàntica profunda i la detecció de desinformació. Encara que comparteix la base tècnica del Transformer, el seu disseny incorpora tècniques com Retrieval-Augmented Generation (RAG) (Lewis et al. 2020), que vincula el model a bases de dades externes, per contrastar afirmacions en temps real. A més, el seu entrenament multilingüe especialitzat (espanyol, xinès, anglès) i la seva eficiència computacional el posicionen com una eina escalable per processar fluxos massius de dades.

La introducció de DeepSeek com una alternativa de codi obert i gratuïta a ChatGPT marca un punt d'inflexió en el camp de la IA, oferint noves possibilitats a la investigació científica. Si bé ambdós models comparteixen fortaleses significatives en la millora de l'eficiència i la democratització de l'accés, les diferències en la seva transparència, cost i accessibilitat tenen implicacions profundes per a la innovació, l'ètica i l'assignació de recursos. La transició de la IA d'una eina assistencial a un col·laborador actiu requereix una adaptació contínua de les pràctiques de recerca (Kayaalp et al., 2025). Ambdós models comparteixen la capacitat de realitzar una anàlisi lingüística profunda, identificant indicadors de manipulació emocional, absència de fonts i contradiccions internes, elements freqüentment associats a la difusió d'informacions falses (Shu et al., 2017).

2.3. Detecció de desinformació amb LLM: avenços i perspectives

Els LLM han transformat el camp de la detecció de desinformació. Els models preentrenats com BERT i RoBERTa ja havien establert una base sòlida per a la classificació de textos (Liu et al., 2019). No obstant això, els LLM actuals, basats en arquitectures Transformer, ofereixen capacitats superiors de comprensió contextual i raonament semàntic, cosa que permet abordar la desinformació de forma més adaptativa (Papageorgiou et al., 2025). Una contribució clau dels LLM és la seva aplicació en escenaris de detecció de notícies falses amb pocs exemples (few-shot fake news detection). Aquesta capacitat és crucial per adaptar-se ràpidament a les noves narratives de desinformació, ja que aquests models poden generalitzar a partir d'un nombre molt reduït d'exemples, superant les limitacions dels mètodes supervisats tradicionals (Patel et al., 2025).

La integració de tècniques d'explicabilitat (Explainable AI) s'ha convertit en un pilar fonamental. Metodologies com LIME (Local Interpretable Model-agnostic Explanations) i SHAP (SHapley Additive exPlanations) permeten desglossar la base racional d'una classificació, identificant paraules o frases clau associades a contingut fals (Patel et al., 2025). Això és essencial tant per a l'auditoria del sistema com per proporcionar retroalimentació útil als verificadors humans. La investigació recent també apunta cap a la detecció multimodal, on els LLM es combinen amb xarxes neuronals per processar informació de text, imatges i vídeos. La construcció de mòduls de fusió de característiques permet identificar inconsistències subtils entre diferents formats, una característica comuna en les campanyes de desinformació modernes (Lu y Yao, 2025).

No obstant això, aquest potencial conviu amb desafiaments importants. Els LLM són propensos a les al·lucinacions —la generació d'afirmacions plausibles però falses— cosa que compromet la seva fiabilitat com a detectors (Pooley, 2024). Així mateix, existeix el risc de canibalisme del coneixement: a mesura que el contingut generat per IA inunda internet, els models futurs podrien entrenar-se amb sortides d'altres models, degradant la qualitat de la seva base de coneixement (Pooley, 2024). La detecció de desinformació camuflada, que imita fonts legítimes, exigeix que els models se centrin en l'anàlisi factual més que en patrons estilístics superficials (García, 2024).

En conjunt, els LLM no són una solució única, sinó eines poderoses l'efectivitat de les quals depèn d'una integració intel·ligent amb mètodes de verificació externa i una comprensió crítica de les seves limitacions. ChatGPT i DeepSeek ofereixen diverses capaci-

tats prometedores, però la que afecta el present treball consisteix en la seva capacitat de «llegir» la web per processar i «raonar» el seu contingut. D'aquesta manera, l'usuari pot crear ràpidament una eina de scraping sense necessitat de comptar amb coneixement de codificació i beneficiant-se del model de llenguatge per a la seva anàlisi. A continuació, es detallen els passos seguits en aquest experiment.

3. Objectius

L'objectiu primari d'aquest article aborda la presència dels grans editors científics en aquests canals i la possibilitat de detectar els possibles fake channels amb ChatGPT i DeepSeek. Per a això proposem respondre a les següents preguntes de recerca:

1. Existeixen fake channels a Telegram dels editors acadèmics més importants?
2. Podrien servir ChatGPT i DeepSeek per identificar aquests canals?
3. Quines fonts web utilitzen aquests LLM per respondre a la qüestió anterior?
4. Existeix algun biaix en les fonts per països en funció de l'origen de cadascun d'ells?

4. Material i mètodes

El present estudi adopta un disseny d'estudi de cas múltiple, seguint els lineaments metodològics establerts per Yin (2018). Segons aquest autor, l'estudi de cas constitueix una investigació empírica que examina un fenomen contemporani dins del seu context real, sent especialment apropiat quan les fronteres entre el fenomen i el context no són clarament evidents. L'elecció de casos múltiples (13 editorials i 37 canals associats) respon a la lògica de replicació literal proposada per Yin, on cada cas serveix com una «rèplica» que permet confirmar o refutar els patrons emergents. Addicionalment, com assenyala Eisenhardt (1989), la investigació basada en casos múltiples resulta particularment adequada per a àrees temàtiques noves —com en aquest cas— ja que permet generar teoria empíricament fonamentada a partir de la comparació sistemàtica entre casos. A continuació, es presenta la justificació del cas d'estudi atenent als seus antecedents, el propòsit de la seva selecció i les unitats d'anàlisi.

4.1. Antecedents del cas

La proliferació de canals falsos en plataformes de missatgeria constitueix un fenomen que s'ha intensificat en els últims anys, especialment després de la pandèmia de COVID-19. Telegram presenta característiques estructurals que el fan particularment vulnerable a la suplantació d'identitat. A diferència d'altres plataformes, Telegram s'ha posicionat com un servei amb moderació mínima, cosa que facilita la creació de canals que imiten organitzacions legítimes (Urman i Katz, 2022). L'impacte d'aquestes vulnerabilitats en l'àmbit de la comunicació científica resulta especialment preocupant. Els escàndols de suplantació d'identitat acadèmica han sacsejat el sector editorial en els últims anys, amb casos documentats d'autors falsos, cartes d'acceptació fraudulentes i ús d'intel·ligència artificial per fabricar dades (Stockemer i Reidy, 2024). Aquesta erosió s'estén a l'àmbit científic quan canals no verificats difonen informació sota l'aparença d'editorials acadèmiques reconegudes.

Encara que Telegram ha implementat un sistema de verificació de tercers, la seva adopció per part de les editorials acadèmiques ha estat pràcticament inexistent. Cap de les editorials analitzades en aquest estudi comptava amb canals verificats oficialment a la

plataforma. Aquesta absència de verificació oficial deixa un buit que els actors malintencionats exploten mitjançant l'ús de noms, logos i descripcions que imiten les organitzacions legítimes.

4.2. Propòsit de la selecció

Criteris de selecció d'editorials: Es van seleccionar les 13 principals editorials acadèmiques segons el nombre de títols de revista indexats a SCImago Journal i Country Rank (consulta gener 2025, N > 30k fonts). El criteri de nombre de títols, enfront del de documents publicats, es va adoptar deliberadament per evitar el biaix que introduirien les megajournals i per garantir una representació de l'ecosistema editorial en la seva diversitat. Les editorials seleccionades (Elsevier, Springer, Wiley-Blackwell, Routledge, Oxford University Press, De Gruyter, Brill, Cambridge University Press, IEEE, Hindawi, World Scientific, Nature i Thieme) representen els principals actors del sector i, per tant, els blancs més probables de suplantació. El llistat complet amb les diferents variants es pot trobar a l'Apèndix I.

Justificació de la plataforma Telegram: L'elecció de Telegram com a objecte d'estudi respon a tres factors convergents. En primer lloc, el seu creixent ús per a la difusió de contingut acadèmic i científic, documentat en estudis recents sobre ecosistemes de desinformació (Pietro et al., 2024). En segon lloc, les seves limitacions estructurals de verificació, que contrasten amb plataformes més regulades i generen vulnerabilitats específiques per a la suplantació d'identitat institucional (Herrero i Castro, 2022). En tercer lloc, l'accessibilitat pública del contingut dels canals mitjançant URL directes (t.me/s/[canal]), que permet la seva anàlisi mitjançant LLM amb capacitat de navegació web.

Justificació dels LLM seleccionats: Tal com havíem avançat a la introducció, la selecció de ChatGPT i DeepSeek com a eines d'anàlisi respon a la seva representativitat dins del panorama actual de models de llenguatge. ChatGPT, desenvolupat per OpenAI, manté una posició dominant al mercat occidental amb una quota superior al 80% entre els chatbots d'IA generativa, consolidant-se com el referent comercial per a aplicacions de verificació. DeepSeek representa un contrapunt metodològic rellevant: un model de codi obert desenvolupat sota restriccions d'accés a maquinari avançat que, no obstant això, ha demostrat capacitats competitives. La comparació entre ambdós models permet avaluar si les diferències en arquitectura, entrenament i orientació (comercial vs. codi obert / EUA vs. Xina) influeixen en la capacitat de detecció de canals fraudulents.

4.3. Unitats d'anàlisi

Seguint l'estructura jeràrquica recomanada per Yin (2018) per a estudis de cas múltiple, aquest estudi defineix tres nivells d'anàlisi:

Nivell	Unitat d'anàlisi	N	Descripció
1	Editorials acadèmiques	13	Principals editorials segons SCImago per nombre de títols
2	Canals de Telegram	37	Canals identificats mitjançant cerca pel nom de l'editorial
3	Avaluacions de LLM	74	ChatGPT (37) + DeepSeek (37)

Taula 1. Unitats d'anàlisi

Variables recopilades per a cada canal: nom del canal, compte (@usuari), editorial a la qual s'associa, nombre de seguidors, text del perfil, classificació emesa (REAL/FAKE/UNK) per cada LLM, fonts web citades per cada LLM, domini geogràfic de les fonts, i finalment etiquetat de referència (ground truth) establert pels investigadors.

Delimitació temporal: Les avaluacions es van realitzar utilitzant les versions de ChatGPT-V4o i DeepSeek-V3, entre l'1 i el 15 de febrer de 2025, ambdues amb la funció de cerca web activa.

Delimitació d'abast: Es van incloure únicament canals públics el nom, descripció o contingut dels quals pogués generar confusió amb les editorials oficials. Es van excloure grups privats (no accessibles públicament) i canals clarament paròdics o satírics sense intenció de suplantació.

4.4. Procediment de recopilació de dades

El procediment de recopilació va seguir un protocol estandarditzat per garantir la replicabilitat:

Identificació de canals: Per a cadascuna de les 13 editorials, es va realitzar una cerca a Telegram utilitzant el nom de l'editorial. El sistema de cerca de Telegram mostra un màxim de 10 resultats per consulta. Es van conservar tots els canals que poguessin ser confosos a simple vista amb l'editorial oficial.

Registre de dades: Per a cada canal identificat es va registrar el nom, el compte (@usuari), el nombre de seguidors i el text del perfil en un dataset estructurat.

Avaluació mitjançant LLM: Es va llançar a ChatGPT i DeepSeek el següent prompt estandarditzat per a cada canal:

> You are an expert in Telegram application channels. Could you indicate if the channel @[nom_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/[nom_canal]. What facts do you base your assessment on to determine if it is fake or not?

El prompt es va formular en anglès per maximitzar la capacitat d'ambdós models d'accedir a fonts web internacionals. Cada avaluació es va realitzar en un xat nou, sense activar el pensament profund (deep thinking) i amb l'opció de cerca web activa.

Registre de respostes: Es van documentar les respostes completes de cada model, incloent la classificació emesa i els dominis web citats com a fonts de verificació.

Verificació manual: Els investigadors van establir una classificació ground truth mitjançant verificació manual de cada canal, consultant les pàgines web oficials de les editorials, els seus perfils verificats en altres xarxes socials i el contingut històric dels canals.

5. Resultats

En primer lloc, tenim a la figura 1 un rànquing de les editorials amb la composició de títols de revistes per països. Trobem un fort predomini dels Estats Units, el Regne Unit i els Països Baixos, en general, i d'alguns països en editors concrets com Alemanya a Springer, de Gruyter i Thieme o Egipte a Hindawi.

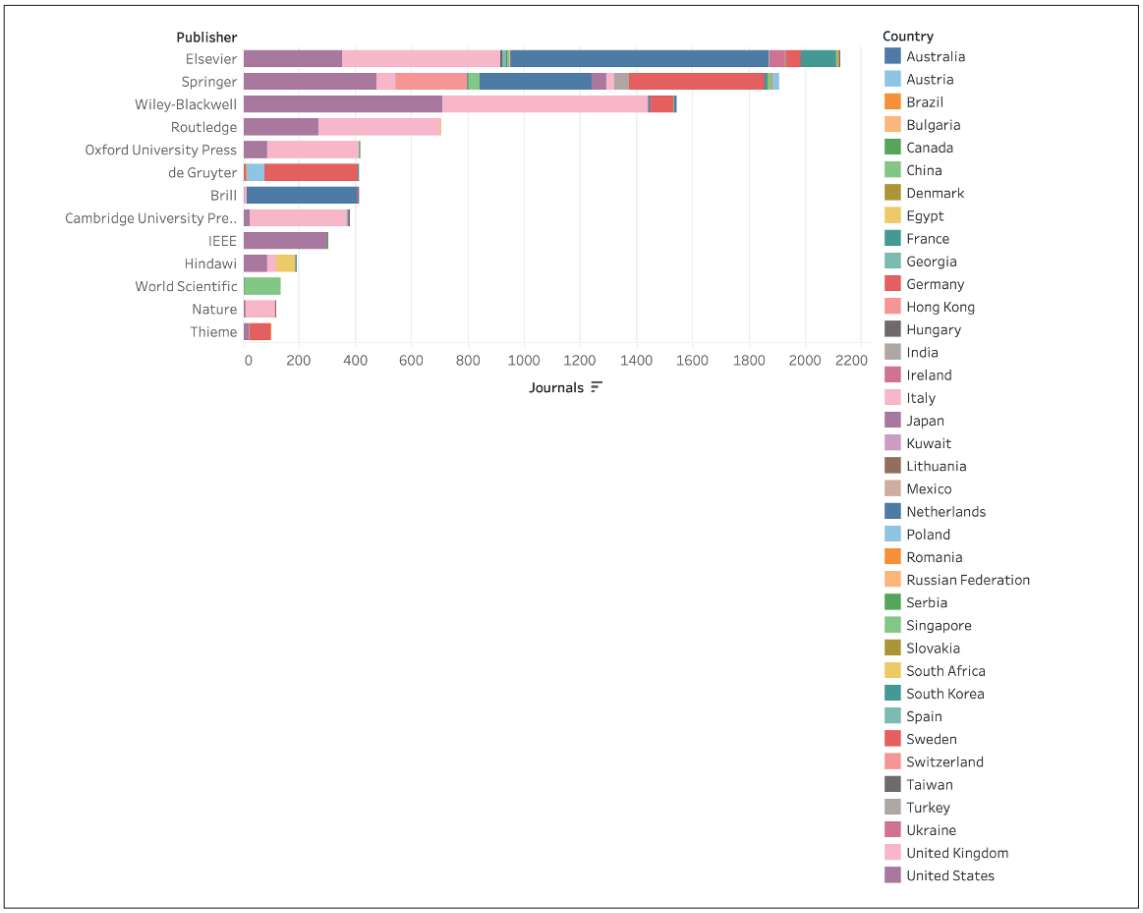

Figura 1.Volum de títols per països en cada editor

5.1. Valoració de les respostes

Cal assenyalar que en cap dels 37 canals analitzats apareix la marca de verificació que realitza Telegram dels canals oficials; en molt pocs d'ells s'afegeix una URL, que és un element essencial en la verificació, observant-se que la informació publicada als canals és poc homogènia. Ocorre quelcom semblant amb el nombre de seguidors i el nombre de publicacions de cada canal, que resulten molt dispars. En l'anàlisi manual dels canals es va concloure que només 8 de 37 eren canals reals i que estaven efectivament vinculats amb les editorials, un 21,62% del total. A la figura 2 veiem aquests resultats.

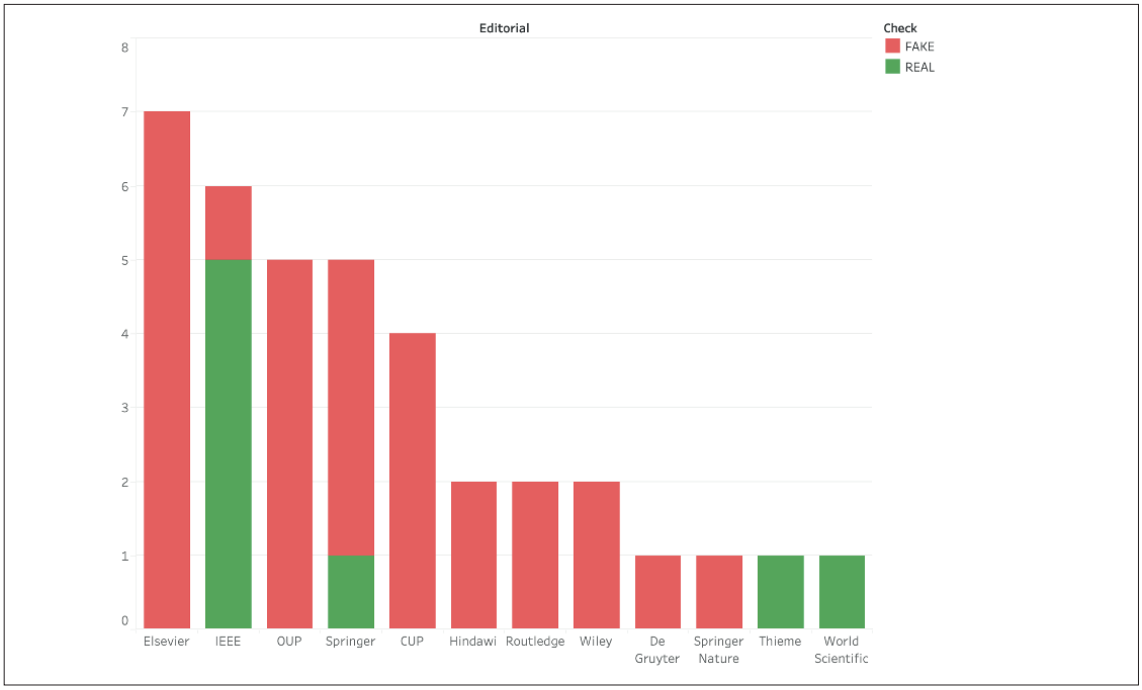

Figura 2. Relació de canals veritables i falsos per editorial

L'anàlisi del contingut dels canals de Telegram, fet pels models, revela conclusions detallades sobre el que publiquen aquests canals, cosa que és fonamental per determinar si són reals o falsos. Pel que fa a l'anàlisi i detecció de continguts fraudulents, diversos canals són considerats com a FAKE basant-se en el tipus de serveis i el llenguatge utilitzat, que és inconsistent amb les pràctiques de les editorials de prestigi, com anuncis repetits per a l'enviament d'articles, tarifes de publicació i terminis d'enviament ràpids. Crida l'atenció que apareguin afirmacions inusuals, com temps de revisió i correccions de proves gratuïtes i indexació en revistes per una tarifa, pràctiques associades a revistes depredadores.

Els models mostren encert en detectar publicacions amb continguts fraudulents i localitzar afirmacions contradictòries amb les polítiques de comercialització de les pròpies editorials, apareixent llibres en diferents formats sense proporcionar informació clara sobre drets d'autor o llicències. Els models també detecten presències incongruents, com la distribució de materials de Cambridge University Press, malgrat portar el nom d'Oxford University Press.

Pel que fa al contingut consistent amb afiliació oficial o legítima, en alguns canals, principalment aquells afiliats a IEEE, localitza continguts que recolzen (o almenys no refuten) la seva legitimitat local o temàtica. L'anàlisi del contingut dels canals demostra que la majoria dels canals FAKE utilitzen el contingut (llibres gratuïts, promeses de publicació ràpida) com a esquer per atreure usuaris, mentre que els canals classificats com a REALS o legítimament afiliats (principalment IEEE) es circumscriuen a continguts estrictament professionals, tècnics i vinculats a plataformes oficials de l'organització. Els models també tenen en compte, encara que en menor mesura, el fet de la inactivitat dels canals o les escasses publicacions d'alguns.

En la revisió manual i en les respostes dels models, en els canals FAKE s'observa que hi ha una intencionalitat de suplantació, utilitzant diferents derivacions en el nom del canal i la seva identificació, amb clara intenció de suplantació de l'editorial, en alguns casos per a la difusió gratuïta de publicacions que es comercialitzen en webs oficials. Pel que fa als usuaris i el nombre de publicacions, són nombrosos els canals que tenen escasses publicacions i que es troben inactius des de fa anys.

També s'observa que la majoria dels corresponents a IEEE són reals. Això es deu al fet que IEEE és una associació de professionals a més de ser una editorial; aquests canals dediquen les seves publicacions a qüestions diferents, i encara que hi ha referències i enllaços a publicacions, la seva creació i el seu ús no responen exclusivament a finalitats editorials.

De la mateixa manera, existeix un reduït grup de petites editorials que utilitzen el canal com un element de difusió d'informació sobre les seves publicacions i activitats diverses, però el fet que no hagin gestionat amb Telegram el seu check de veracitat podria ser un indicatiu que el canal no és una prioritat en les seves polítiques de difusió i comercialització.

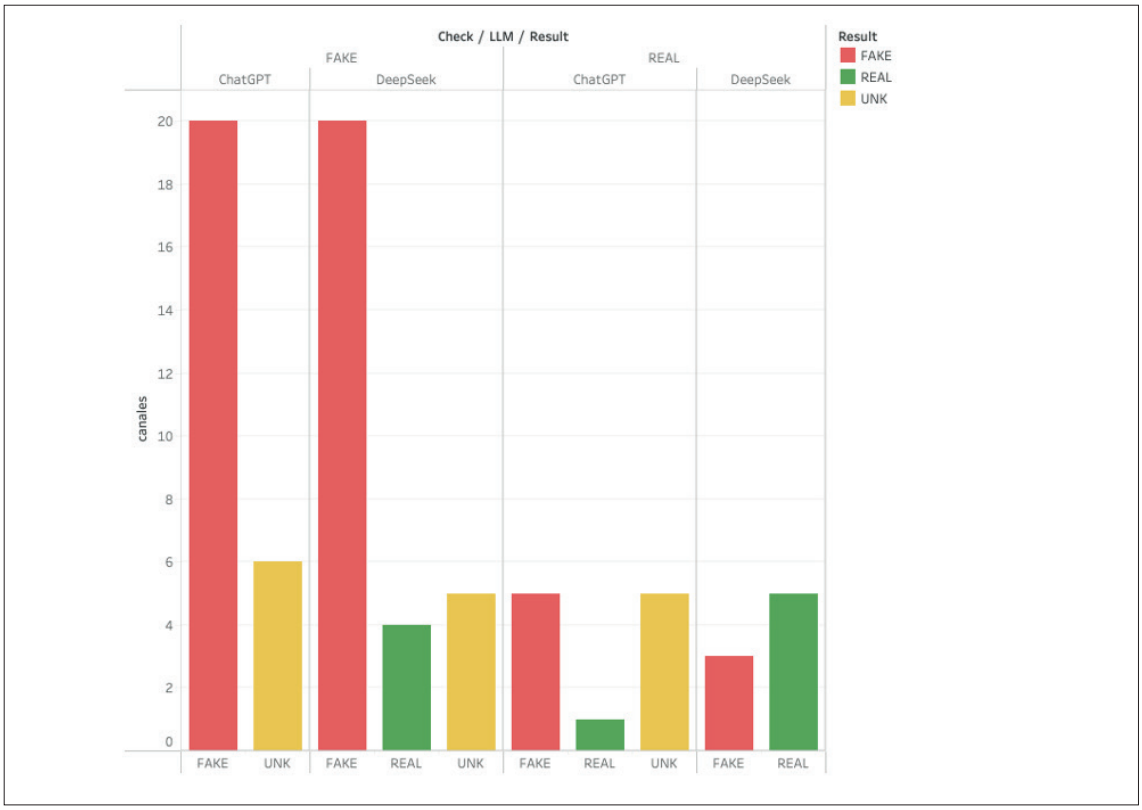

Figura 3. Relació de canals veritables i falsos per consideració dels models.

A la figura 3 es mostra un alt nivell d'encert d'ambdós models en la determinació de la falsedat dels canals FAKE. En el cas de ChatGPT no considera real cap dels falsos, considerant com a dubtosos un grup semblant al considerat per DeepSeek com a tals. Destaca el grup de FAKE que DeepSeek considera reals, transllluint-se que l'enviament des dels continguts a dominis legítims i d'institucions oficials poden haver estat causants de la consideració com a reals.

En el cas de la resposta davant els canals reals s'observa que DeepSeek reparteix les seves respostes entre reals i falsos, no considerant dubtós cap, mentre que ChatGPT limita el nombre dels que considera reals a un, considerant dubtosos o falsos un grup semblant. Això mostra que els models tendeixen a fallar en la identificació de canals reals a causa de la dependència de criteris de verificació global (insígnia blava i enllaços corporatius), ignorant la legitimitat del contingut o l'afiliació regional.

5.2. Llocs referenciats

Al dataset podem trobar els rànquings dels dominis més utilitzats per cada model. Hi ha un total de 217 llocs web que van ser citats un total de 507 vegades. ChatGPT ha fet un ús més intensiu de les referències, un total de 319 enfront de 188 de DeepSeek. Com que la quantitat de llocs és molt gran, a la taula 2 només veiem els més freqüents de cada LLM.

ChatGPT				Deepseek			
Lloc	**Grup**	**Pais**	**#**	**Lloc**	**Grup**	**Pais**	**#**
t.me	Telegram	EAU	35	t.me	Telegram	EAU	33
facebook.com	Xarxes socials	EUA	15	keepersecurity.com	Ciberseguretat	EUA	11
telemetr.io	Telegram	EUA	9	telegram.org	Telegram	EAU	9
youtube.com	Xarxes socials	EUA	9	infostealers.com	Ciberseguretat	EUA	7
reddit.com	Xarxes socials	EUA	9	lifelock.norton.com	Ciberseguretat	EUA	6

Taula 2. Top 5 de llocs citats

Per poder analitzar tots els llocs els hem agrupat en categories i a més els hem inclòs el país de referència. En alguns casos això ha estat senzill, però en altres no tant. A més de revisar els llocs web, hem usat el servei Whois i fins i tot hem preguntat als mateixos LLM. D'aquesta manera obtenim el rànquing de la taula 3, on tenim cada grup i la freqüència acumulada en la qual apareixen (ChatGPT + DeepSeek). El grup principal és el que constitueix les diferents variants del servei Telegram. És raonable que així sigui, ja que en el propi prompt hem demanat als models que analitzin aquest servei. En segon lloc, tenim els serveis de xarxes socials en general, alguns tan coneguts com Facebook, YouTube o Reddit, però també n'hi ha d'altres menys destacats. Dins d'Editors top incloem els llocs web dels editors objecte d'estudi, i dins d'Editors la resta.

Grup	Freqüència
Telegram	117
Xarxes socials	64
Editors top	52
Ciberseguretat	50
Repositoris	41
Editors	39
Universitats	37
Mitjans	24
Governs	22
Màrqueting digital	17
Black hat acadèmic	16
Cripto	15
Blog educatiu	7
Revistes acadèmiques	3
Fake news	2
Marketplace	1

Taula 3. Freqüència de llocs per grup

És important destacar el grup Ciberseguretat, ja que hi ha hagut una gran quantitat de llocs relacionats amb aquesta temàtica, a priori no esperables, i també va ocórrer el mateix amb tot el relacionat amb el món Cripto. De la resta, potser cal destacar el que hem anomenat Black hat acadèmic i que no és més que llocs web on es facilita l'accés a material amb drets d'autor en una clara (o dissimulada) vulneració d'aquests.

A la figura 4 podem apreciar els biaixos de cada model amb cada grup. Clarament Telegram apareix en ambdós; no obstant això, Ciberseguretat (i Cripto en menor mesura) destaca a DeepSeek. ChatGPT, per la seva part, sembla utilitzar llocs més mainstream com les mateixes xarxes socials, les universitats o llocs del govern. Els Editors, Editors top i Repositoris apareixen referenciats per ambdós models. Cal destacar que el Black hat acadèmic és esmentat per ambdós models, encara que ChatGPT ho fa en major mesura.

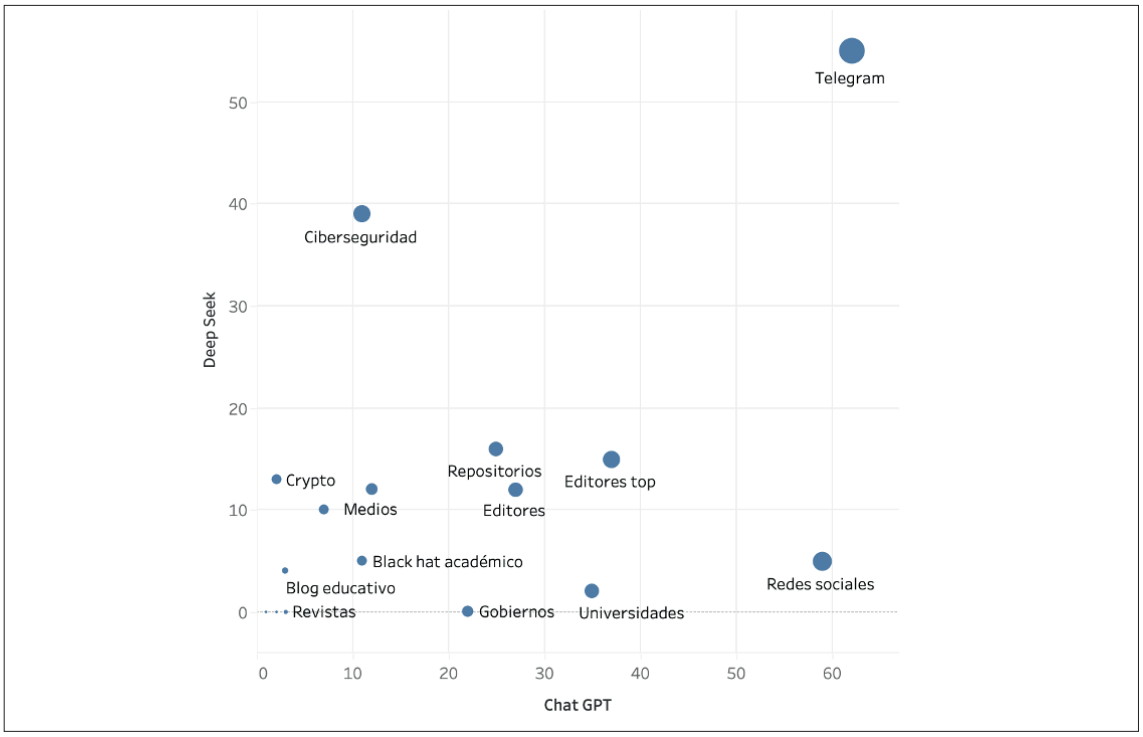

Figura 4 – Comparació de llocs per model

Però el que més crida l'atenció és la procedència dels mitjans per països. En aquest cas no s'aprecia cap biaix destacable. A la figura 5 apareixen representats els països amb dues o més aparicions. El país predominant és els Estats Units, tant per al model estatunidenc com per al xinès. En segon lloc, tenim els Emirats Àrabs Units, ja que és la seu del servei Telegram, i una mica més enrere el Regne Unit i Alemanya. Fins aquí, un domini total de l'espai web occidental. En cinquè i sisè lloc tenim dos actors no occidentals: Rússia i l'Índia. Crida l'atenció que Rússia aparegui esbiaixada cap a ChatGPT mentre que l'Índia ho fa cap a DeepSeek, on especialment l'ha utilitzat dins de la temàtica de la ciberseguretat.

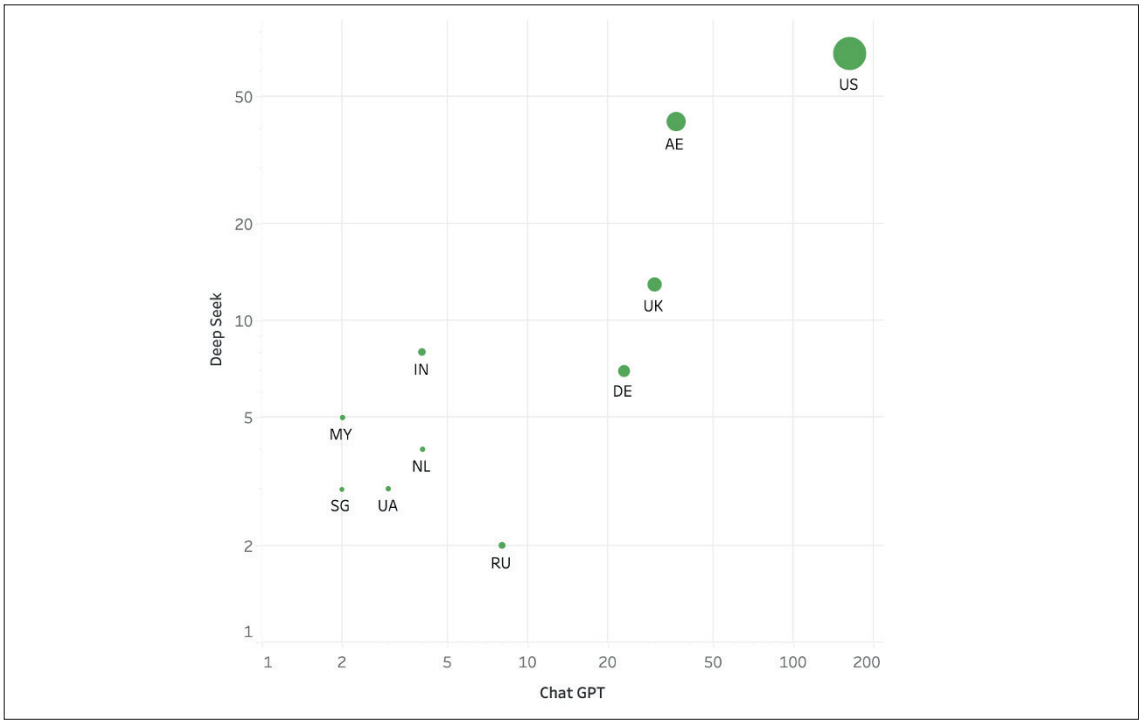

Figura 5. Distribució de països per model

Finalment, hem de destacar que en tot el dataset només hi ha un lloc xinès, cosa realment curiosa.

6. Discussió

La present investigació confirma la proliferació massiva de canals no oficials que suplanten la identitat de les principals editorials acadèmiques a Telegram. Les dades revelen que el 78,38% dels canals analitzats presenten característiques fraudulentes o pràctiques no autoritzades, constituint un ecosistema paral·lel que opera al marge dels canals oficials de comunicació científica. Aquesta troballa és consistent amb investigacions prèvies que han documentat l'existència d'una extensa xarxa de canals falsos i clons a la plataforma (La Morgia et al., 2021, 2023), i estableix que la suplantació d'identitat o l'operació de canals no oficials és la norma, no l'excepció, en aquest ecosistema.

Aquest fenomen no pot dissociar-se del context més ampli de l'hegemonia editorial occidental identificada a la nostra Figura 1, on els Estats Units, el Regne Unit i els Països Baixos concentren la major part de la producció. Aquesta concentració crea inevitablement barreres d'accés econòmiques, idiomàtiques i geopolítiques, que al seu torn generen un brou de cultiu per a mercats informals que prometen eludir aquestes restriccions. L'aparició recurrent a les nostres fonts de categories com «black hat acadèmic» corrobora l'existència d'una infraestructura digital organitzada que sustenta aquesta economia grisa, un fenomen que ha estat identificat també en el context del frau editorial acadèmic (Stockemer i Reidy, 2024).

Pel que fa als mecanismes de frau identificats, els canals FAKE utilitzen tàctiques específiques inconsistents amb la comunicació acadèmica legítima. En primer lloc, la promoció fraudulenta, caracteritzada per un llenguatge informal i promocional, amb afirmacions inusuals com terminis de publicació extremadament ràpids (2-3 setmanes) i l'esment de tarifes elevades (1.200 USD), que s'alineen amb les característiques de revistes depredadores o estafes. En segon lloc, la distribució no autoritzada i pirateria: els canals relacionats amb editorials de llibres (OUP, Springer) ofereixen accés gratuït a materials que l'editor normalment comercialitza, cosa que resulta atípica per a un editor legítim. Aquests patrons coincideixen amb els descrits per Herasimenka et al. (2023) en la seva anàlisi de plataformes amb moderació mínima, on la facilitat de creació de canals i l'absència de verificació preventiva faciliten la proliferació de continguts enganyosos. La característica més comuna dels canals no oficials és l'absència del segell de verificació blau de Telegram, que editorials de la talla de les que estem estudiant haurien d'ostentar si el canal fos oficial.

La gran presència de canals FAKE, que sovint combinen la distribució de contingut pirata amb serveis de publicació sospitosos, representa un risc significatiu per a la integritat acadèmica i la marca dels editors. Aquesta troballa ve a ratificar que la possibilitat de generar qualsevol canal públic, amb qualsevol denominació, en estar prevista la verificació només a posteriori i únicament si la corporació o entitat la sol·licita, obre un espai notable a les possibilitats de frau. Resulta cridaner que les editorials científiques no semblin mostrar un interès especial per aquesta via que, ben explotada i correctament verificada, podria constituir un mitjà de difusió robust per a les mateixes. Aquesta desatenció institucional contrasta amb la creixent rellevància de Telegram com a plataforma de difusió informativa (Willaert, 2023; Sánchez i Martos, 2020).

Respecte a la idoneïtat dels models d'LLM com a eines d'anàlisi, els nostres resultats presenten un panorama matisat. La significativa coincidència entre ambdós models en la identificació de canals clarament fraudulents suggereix que aquestes eines poden efectivament detectar patrons comuns de suplantació, cosa que és consistent amb

investigacions recents que demostren la capacitat dels LLM per identificar continguts enganyosos mitjançant anàlisi contextual (Chen i Shu, 2024; Papageorgiou et al., 2025). No obstant això, les notables discrepàncies en la verificació dels canals reals revelen limitacions estructurals profundes. DeepSeek demostra una tendència marcada cap al que podríem denominar un «criteri contextual», mostrant major disposició a validar canals que presenten contingut coherent i col·laboracions locals, fins i tot davant l'absència de verificació formal. En contrast, ChatGPT opera sota un paradigma de «verificació estricta», requerint evidències explícites d'afiliació institucional i mostrant una cautela considerablement major.

Aquesta divergència en els criteris d'avaluació té implicacions importants. Els models poden ser enganyats per l'aparença professional del contingut o per l'ús de termes tècnics (com Scopus, Nature, etc.), sense poder validar els indicadors crítics d'autenticitat. Aquesta limitació connecta amb les advertències sobre les al·lucinacions dels LLM —la generació d'afirmacions plausibles però falses— que comprometen la seva fiabilitat com a detectors autònoms (Bender et al., 2021; Pooley, 2024). Els resultats reforcen la necessitat de considerar aquestes eines com a suport complementari al judici expert, no com a substituts del mateix, tal com suggereixen Patel et al. (2025) en la seva proposta de sistemes de detecció de desinformació.

L'examen de les fonts web consultades pels models revela arquitectures de coneixement notablement diferents. Més enllà del predomini esperable de dominis associats a Telegram, emergeix amb força inusitada l'àmbit de la ciberseguretat en el cas de DeepSeek, mentre que ChatGPT mostra preferència per fonts institucionals i xarxes socials convencionals. Aquesta disparitat suggereix que cada model construeix el seu marc interpretatiu sobre agents de cerca web substancialment diferents, cosa que necessàriament condiciona les seves avaluacions. DeepSeek, en recolzar-se en recursos especialitzats en seguretat digital, sembla desenvolupar una sensibilitat aguda cap a patrons de frau i suplantació, un enfocament que recorda les tècniques de detecció basades en anàlisi de patrons descrites per Shu et al. (2017). ChatGPT, amb la seva base més àmplia i convencional, prioritza la verificació institucional formal. Ambdós enfocaments presenten avantatges complementaris, però la seva desconnexió explica en gran mesura les discrepàncies observades en les classificacions.

El predomini absolut de fonts occidentals en ambdós models reflecteix l'hegemonia encara vigent dels anomenats països desenvolupats en l'arquitectura global d'internet, un biaix que ha estat documentat en estudis previs sobre la distribució geogràfica del coneixement en línia (Allcott i Gentzkow, 2017). No obstant això, les afinitats diferencials —DeepSeek amb fonts índies i del sud-est asiàtic, ChatGPT amb russes— suggereixen que els agents de cerca de fonts no estan completament predeterminats. El cas particularment eloqüent de la gairebé absència de fonts xineses, malgrat l'origen de DeepSeek, mereix investigació addicional. Aquesta troballa connecta amb les observacions d'Urman i Katz (2022) sobre com les estructures de xarxa a Telegram poden replicar divisions ideològiques i geogràfiques més àmplies.

Finalment, els nostres resultats tenen implicacions per a la comprensió del fenomen més ampli del desordre informatiu descrit per Wardle i Derakhshan (2017). Els canals falsos de Telegram que suplanten editorials acadèmiques no constitueixen únicament un problema de propietat intel·lectual o frau comercial, sinó que erosionen la confiança en les institucions de comunicació científica. En un context on la desinformació científica ha demostrat tenir conseqüències greus per a la salut pública i el debat democràtic (Díez-Garrido et al., 2021; Prieto-Campo et al., 2024), la proliferació d'aquests canals representa una amenaça estructural que requereix respostes coordinades tant de les plataformes tecnològiques com de les mateixes editorials acadèmiques.

7. Conclusions

Aquest treball posa de manifest que, a l'ecosistema de Telegram, la presència de les grans editorials científiques està profundament distorsionada per una majoria aclaparadora de canals no oficials que suplanten la seva identitat. A partir de l'anàlisi de 37 canals associats a 13 editorials líders, es va constatar que el 78,38% dels canals són fraudulents o manquen de vincles verificables amb l'editorial corresponent, mentre que només un 21,62% pot considerar-se real o legítimament afiliat. Aquesta situació configura un entorn estructuralment propens a la desinformació i a la vulneració de drets de propietat intel·lectual.

En relació amb la primera pregunta de recerca, els nostres resultats confirmen l'existència d'una xarxa extensa i organitzada de canals no oficials que s'apropien de noms, logotips i descripcions pròpies de les editorials acadèmiques. Aquests canals combinen la distribució no autoritzada de continguts (especialment llibres) amb ofertes de serveis editorials de dubtosa legitimitat, alineats amb pràctiques pròpies de revistes depredadores. L'absència generalitzada de verificació oficial de Telegram i la limitada presència activa de les editorials a la plataforma afavoreixen que aquests canals fraudulents es converteixin, de facto, en referents visibles per a molts usuaris.

Quant a la segona pregunta de recerca, els resultats mostren que ChatGPT i DeepSeek són eines útils per al mapatge i la caracterització inicial d'aquest ecosistema fraudulent, especialment en la identificació de canals clarament FAKE. Ambdós models coincideixen amb freqüència en la detecció de patrons de frau. No obstant això, presenten limitacions importants en la validació de canals reals, on tendeixen a sobrereaccionar davant l'absència de senyals institucionals forts (canals verificats, enllaços web corporatius) i a infravalorar continguts legítims amb visibilitat local o regional. La diferència entre l'enfocament més «contextual» de DeepSeek i la major exigència de «verificació formal» de ChatGPT suggereix que aquests models operen sota lògiques d'avaluació parcialment divergents, cosa que reforça la necessitat d'utilitzar LLM com a suport complementari, i no com a substitut, del judici expert humà.

Amb relació a la tercera i quarta pregunta, la temàtica de les fonts consultades revela que la construcció de coneixement en aquests sistemes està mediada per arquitectures documentals específiques que reflecteixen prioritats temàtiques i alineaments diferenciats. El biaix cap a la ciberseguretat de DeepSeek i l'«institucionalisme» (universitats, governs, editors) de ChatGPT no semblen constituir meres diferències tècniques, sinó visions complementàries d'un fenomen complex. No obstant això, aquesta diferència no es verifica clarament quan mirem l'origen de les fonts. Existeix un biaix cap a fonts occidentals, cosa que sembla indicar que els entrenaments d'ambdós models s'han basat en conjunts de dades no gaire diferents. La presència de fonts xineses és gairebé nul·la (0,2%), fins i tot en el cas de DeepSeek, un model que per moments dona la impressió de ser indi o del sud-est asiàtic (Malàisia i Singapur). En conjunt, els resultats suggereixen que les diferències entre ambdós LLM s'expliquen més per la lògica dels seus agents de cerca i la priorització temàtica de les seves fonts que pel país d'origen de cada model.

Com a línies de treball futures, aquest estudi obre diverses vies d'interès. En el pla metodològic, resulta pertinent incorporar altres LLM rellevants. D'entre ells destaquen actualment, segons el nostre criteri, Claude (Anthropic), Gemini (Google), Qwen (Alibaba), el controvertit Grok (X) i el recent arribat Kimi K2. Amb ells podrem avaluar si les diferències observades en aquest treball es mantenen en un espectre més ampli de models. En el pla aplicat, seria valuós estendre l'enfocament a altres tipus de desinformació científica i política presents a Telegram, incloent la detecció de fake news i narratives conspiratives. La progressiva integració de capacitats d'anàlisi massiva de dades per part dels LLM ofereix una oportunitat per dissenyar sistemes híbrids en els

quals l'escala computacional i el criteri humà es reforcin mútuament en la protecció de la integritat de la comunicació científica.

Agraïments

Aquest article ha rebut finançament del programa Projectes de Generació del Coneixement 2023 del Ministeri de Ciència, Innovació i Universitats d'Espanya pel projecte «Intel·ligència Artificial a Europa: Auge o declivi?» (PID2023-149646NB-I00).

8. Bibliografia

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. Apèndix I

Editorials amb les seves variants de denominació

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. Apèndix II

Conclusions de model per canal i revisió 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
@elsevieeducatiRealn	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 = fals / UNK = dubtós / REAL = real