



Dra. Daniela Musicco-Nombela
Profesora Titular. Universidad Francisco de Vitoria

@ daniela.musicco@ufv.es

ID 0000-0001-7325-9978

■ Páginas / Pages
De la 13 a la 26
■ ISSN: 1885-365X

Censura digital. Verdad y mentira en la era de la inteligencia artificial. Comunicar la verdad

Digital censorship. Truth and lies in the artificial intelligence era
Communicating the truth

RESUMEN:

Este artículo analiza la censura digital como una arquitectura de visibilidad que reordena la frontera entre verdad y mentira en la era de la inteligencia artificial, desplazando los controles editoriales ex ante hacia mecanismos algorítmicos y procedimentales (selección de datos, ranking, etiquetado, shadow banning) que afectan al pluralismo y a la formación del juicio público, incluso bajo marcos garantistas como el art. 20 CE y la DSA. Los objetivos planteados han sido: 1) Conceptualizar un modelo multicapa de censura digital (estructural, operativa y forense); 2) evaluar sus efectos epistémicos sobre la veracidad pública (accesibilidad, persistencia, prioridad); 3) contrastar el papel de las agencias verificadoras en la balanza entre corrección cívica y silenciamiento; 4) proponer una respuesta educativa experiencial centrada en «buscar la verdad» y comunicar con rigor. Para lograr estos objetivos se ha aplicado la metodología de revisión crítica interdisciplinar (filosofía, derecho, comunicación) y análisis de casos recientes en plataformas (cambios de políticas, etiquetado, despriorización) triangulados con normativa europea; diseño-acción docente mediante el proyecto NCM-BV (2025), que implementa talleres de auditoría de comunicación, trazabilidad de evidencias, creación de narrativas, simulaciones de crisis y ejercicios de puesta en escena para medir competencias de comunicación y ética informativa. Entre los resultados obtenidos, el mapa multicapa explica cómo la verdad pública se debilita cuando deja de circular: la capa estructural determina qué entra en el sistema; la operativa modula la interacción (filtros, bloqueos, verificaciones); la forense reescribe o desprioriza lo ya publicado. En el aula-laboratorio, los estudiantes incrementan la detección de sesgos y la documentación de procesos (bitácoras, capturas con metadatos, cadenas de custodia), mejorando la calidad argumentativa y la resistencia a la manipulación. De las principales conclusiones se deduce que la censura digital opera menos como prohibición y más como gestión de visibilidad; sostener un ecosistema veraz exige transparencia procedimental, límites claros a la intervención verificadora y una didáctica experiencial que forme en mirada crítica, trazabilidad y responsabilidad profesional para educar en libertad y comunicar con evidencia.

KEYWORDS:

Comunicación verdad, buscar verdad, educar verdad, censura digital, agencias verificadoras, inteligencia artificial, libertad de expresión digital.

RESUMEN:

This article analyzes digital censorship as an architecture of visibility that reorders the boundary between truth and falsehood in the age of AI, shifting ex-ante editorial controls toward algorithmic and procedural mechanisms (data selection, ranking, labeling, shadow banning) that affect pluralism and the formation of public judgment, even within rights-protecting frameworks such as Article 20 of the Spanish Constitution and the DSA. The objectives pursued were: (1) conceptualizing a multilayer model of digital censorship (structural, operational, and forensic); (2) assessing its epistemic effects on public veracity (accessibility, persistence, priority); (3) contrasting the role of fact-checking agencies in the balance between civic correction and silencing; and (4) proposing an experiential educational response centered on «seeking truth» and communicating with rigor. To achieve these goals, the methodology applied consisted of an interdisciplinary critical review (philosophy, law, communication) and analysis of recent platform cases (policy changes, labeling, deprioritization) triangulated with European regulation; and a design-based teaching intervention through the NCNM-BV (2025) project, which implements workshops on communication auditing, evidence traceability, narrative construction, crisis simulations, and performance-based exercises to evaluate communication competencies and ethical literacy. Among the results, the multilayer map explains how public truth weakens when it ceases to circulate: the structural layer determines what enters the system; the operational layer modulates interaction (filters, blocks, verifications); and the forensic layer rewrites or deprioritizes already-published content. In the classroom-laboratory, students increase their ability to detect bias and document processes (logbooks, metadata-preserved captures, chains of custody), improving argumentative quality and resistance to manipulation. The main conclusions indicate that digital censorship operates less as prohibition and more as visibility management; sustaining a truthful ecosystem requires procedural transparency, clear limits on the intervention of verifying bodies, and an experiential pedagogy that cultivates critical perception, traceability, and professional responsibility to educate in freedom and communicate with evidence.

KEYWORDS:

Communication truth, seeking truth, educating for truth, digital censorship, fact-checking agencies, artificial intelligence, digital freedom of expression.



1. Introducción

Desde una perspectiva histórica y filosófica, la preocupación por la verdad y la mentira ha acompañado al pensamiento occidental desde sus orígenes. Ya en la Antigüedad tardía, san Agustín advertía sobre la fragilidad moral de un alma que se aparta de la verdad para confiar en la falsedad y alimentar lo efímero, cuestionando así la integridad de quien acepta la mentira como fundamento de su conocimiento y acción: «Un alma que suspira por tales patrañas, ¿no se prostituye apartándose de ti, confía en la mentira y alimenta el viento?» (Agustín, 354-430)¹. Esta reflexión adquiere hoy una especial relevancia en un contexto marcado por la expansión de las tecnologías digitales y, de manera particular, por el desarrollo acelerado de la inteligencia artificial y la consiguiente multiplicación de las *fake news*, los bulos, la manipulación, las mentiras, en definitiva, las patrañas.

En el ámbito jurídico contemporáneo, la defensa de la libertad de expresión ha sido considerada uno de los pilares fundamentales de los sistemas democráticos. En España, la Constitución de 1978 consagró de forma explícita esta garantía en su artículo 20, al reconocer y proteger el derecho a expresar y difundir libremente pensamientos, ideas y opiniones por cualquier medio de reproducción, así como al establecer de manera inequívoca la prohibición de cualquier forma de censura previa:

1. Se reconocen y protegen los derechos: a) A expresar y difundir libremente los pensamientos, ideas y opiniones mediante la palabra, el escrito o cualquier otro medio

1/ Agustín de Hipona (2007). *Confesiones* (ed. y trad. en español). Madrid: Biblioteca de Autores Cristianos (obra original escrita entre 397-400).



Imagen 1. Cuadro de Masini *Testa bendata*. Fuente: Artmajeur (2026).

de reproducción. [...] 2. El ejercicio de estos derechos no puede restringirse mediante ningún tipo de censura previa (Constitución Española, 1978, art. 20.1.a y 20.2)².

Este blindaje normativo respondía a la voluntad de consolidar un marco de libertades tras décadas de restricción y control informativo. Una forma jurídica de asegurar para el presente y el futuro la libertad de expresión en la comunicación. Sin embargo, apenas cinco décadas después, la vigencia efectiva de estos principios se ve sometida a nuevas tensiones. La emergencia de la era digital y el uso creciente de sistemas basados en inteligencia artificial han reabierto el debate sobre los límites de la libertad de expresión, la regulación de los contenidos y las nuevas formas de censura, muchas de ellas de carácter indirecto, algorítmico o tecnológico. En este escenario, la frontera entre la protección frente a la desinformación y la restricción ilegítima de la libertad comunicativa se vuelve cada vez más difusa, planteando interrogantes fundamentales sobre la relación entre verdad, poder y tecnología en las democracias contemporáneas.

En un cuadro del pintor Tiziano Masini, *Testa bendata*³, se muestra la cabeza vendada de una persona, tapada hasta la nariz, las orejas y los ojos están envueltos, solo le queda libre la boca, abierta en un grito sordo: es el rostro de alguien privado de los sentidos fundamentales para la comunicación y el entendimiento, una manifestación de desesperación que, paradójicamente, nadie puede escuchar. La imagen evidencia el control sobre la percepción y la comunicación, donde no solo se borra la palabra, sino también la capacidad de percibir la realidad. A pesar de todo, es un grito de resistencia que, aunque silenciado o silencioso, sigue existiendo como un acto de lucha contra la anulación impuesta y la censura.

2/ Constitución Española (1978). Boletín Oficial del Estado, n.º 311, de 29 de diciembre de 1978. <https://www.boe.es>. Título I. De los derechos y deberes fundamentales. Capítulo segundo. Derechos y libertades. Sección 1.ª De los derechos fundamentales y de las libertades públicas. Artículo 20. Congreso de los diputados. <https://app.congreso.es/consti/constitucion/indice/titulos/articulos.jsp?ini=20&tipo=2>

3/ Tiziano Masini (2015). *Testa bendata*. <https://www.artmajeur.com/tiziano-masini/es/obras-de-arte/11353685/testa-bendata>



2. Nueva censura digital

Con la llegada de internet, las redes sociales y las plataformas digitales, las posibilidades de acceso, difusión y producción de información y comunicación han tenido un crecimiento exponencial; con esta proliferación de canales y espacio ampliado de comunicación, aparentemente, la posibilidad de libertad de comunicación ha alcanzado niveles sin precedentes, pero al mismo tiempo han surgido nuevos mecanismos de control, limitación de contenidos e intentos de descalificación y censura. La inteligencia artificial hace uso de modelos avanzados para filtrar y controlar la información; en algunos países, como China, se utiliza para identificar y limitar contenidos políticamente sensibles, impidiendo la difusión de opiniones críticas. Por ejemplo, el *chatbot* DeepSeek evita responder preguntas sobre temas controvertidos como la represión de Tiananmen o la situación de los uigures. Esta forma de censura no se limita solo a la política y a los países declarados como dictaduras, también afecta a la medicina, la farmacéutica, la educación, la filosofía, la investigación o el arte; las plataformas occidentales aplican filtros basados en inteligencia artificial para moderar contenidos, a menudo con algoritmos que pueden resultar opacos o estar influenciados por intereses económicos. La moderación de contenidos ocurre de forma flagrante en las plataformas sociales de todas las democracias a través de la aplicación de políticas para filtrar información considerada dañina, también a través del controvertido uso de las agencias de verificación⁴. Las agencias de verificación, apelando a una supuesta mediación de comprobación de la veracidad de los contenidos, ejercen el poder de censurar y aplicar la etiqueta/sello de bulo o *fake* a aquella información que se define como falsa. Estas agencias u otras entidades públicas ejerciendo una supuesta persecución, por ejemplo, del discurso de odio o la desinformación, pueden no solo censurar y etiquetar, sino hacer desaparecer ciertos contenidos, cerrar cuentas con condenas temporales o permanentes. El ejercicio de estas agencias de verificación y la potestad de instituciones públicas es amparado en nuevas leyes como la Ley de Servicios Digitales Europea⁵. Esta normativa busca regular supuestamente el funcionamiento de las plataformas digitales, mejorar la transparencia, proteger los derechos de los usuarios y combatir la desinformación en línea⁶. La ley exige que las plataformas digitales colaboren con verificadores de hechos acreditados para teóricamente reducir la difusión de información falsa. Estas medidas pueden ser arbitrarias y restringir la libertad de comunicación⁷. Algunas plataformas limitan la cantidad de veces que un usuario puede compartir un contenido. Por ejemplo, en WhatsApp se implementó una restricción que impide reenviar mensajes virales a más de cinco contactos a la vez; esto se hizo para evitar la propagación de desinformación, pero también puede afectar la difusión de ciertos mensajes



4/ Musicco-Nombela, D., Martínez-Fresneda, H., Zazo, L. (2024) La universidad como lugar para fortalecer... Tirant lo Blanch.

5/ La Ley de Servicios Digitales (DSA) (Digital Services Act) fue propuesta el 15 de diciembre de 2020, aprobada por el Parlamento Europeo el 5 de julio de 2022 y recibió la aprobación final del Consejo Europeo el 4 de octubre de 2022. Se publicó en el Diario Oficial el 19 de octubre de 2022 y comenzó a aplicarse a las principales plataformas el 25 de agosto de 2023, entrando en vigor completamente el 17 de febrero de 2024.

6/ Comisión Europea (2025). Una Europa Adaptada a la Era Digital: nuevas normas para las plataformas en línea. https://commission.europa.eu/strategy-and-policy/priorities-2019-2024/europe-fit-digital-age/digital-services-act/europe-fit-digital-age-new-online-rules-platforms_es

7/ Se reconoce a ciertas organizaciones como fuentes confiables para señalar contenido engañoso o ilegal. Se requiere que las plataformas expliquen cómo funcionan sus sistemas de recomendación y moderación de contenido.

legítimos. Las redes sociales pueden reducir la visibilidad de ciertos contenidos mediante algoritmos que limitan su alcance a través del *shadow banning*, donde una publicación sigue existiendo, pero no aparece en las búsquedas o en los *feeds* de otros usuarios. También se pueden aplicar restricciones a cuentas específicas, reduciendo la frecuencia con la que sus publicaciones aparecen en el algoritmo de recomendación. Las plataformas pueden suspender temporalmente o cerrar cuentas que infringen sus normas; se utiliza como motivación el hecho de que un usuario publique contenido que se considera dañino, que se ha calificado como discurso de odio o desinformación. En algunos casos, las suspensiones pueden ser automáticas, mientras que en otros requieren revisión manual. Además de las redes sociales, los sitios web están también sujetos a las nuevas formas de censura. Este fenómeno, al igual que en las redes sociales, no ocurre solo en países con dictaduras ampliamente reconocidas, sino también en las democracias donde la posibilidad de un verdadero libre acceso a la información se ve dificultado a través de diferentes métodos, algunos inadvertidos; sirva como ejemplo constatar que motores de búsqueda como Google ha comenzado a eliminar progresivamente los dominios nacionales, como Google.es, Google.fr o Google.de, a los que cualquiera tenía libre acceso simplemente tecleando en el buscador, al final del nombre Google, las iniciales de cualquier país, pudiéndose así acceder a los contenidos de noticias o cualquier índole de cada país. Este cambio fue anunciado oficialmente en abril de 2025 en favor de un único dominio global Google.com, simuladamente sin fronteras y bajo la apariencia de una mayor libertad de acceso sin nacionalidades; sin embargo, este cambio realmente se basa en la capacidad del motor de búsqueda para proporcionar resultados según la ubicación geográfica del usuario, sin necesidad de ccTLD (*country code Top-Level Domains*). Esta transición comenzó ya en 2017, cuando Google empezó a determinar la ubicación del usuario directamente desde el dispositivo o la dirección IP, haciendo innecesario el uso de dominios nacionales para la personalización de los resultados, pero impidiendo definitivamente que un usuario con un determinado IP pudiera acceder a información fuera de su zona geográfica de pertenencia IP, incluso viajando. Ahora, al escribir un dominio local como Google.es, se redirige automáticamente a Google.com, pero los resultados de búsqueda han sido limitados a la ubicación del usuario a través de la geolocalización. Además de estos cambios restrictivos en internet, cada vez es más complicado que se ofrezcan resultados en idiomas diferentes al del IP de pertenencia; a menudo, el buscador restringe los contenidos a fuentes del idioma del país en el que reside el IP. Otro mecanismo son los algoritmos de filtrado, herramientas automatizadas que pueden eliminar u ocultar contenido sin suficiente transparencia. Existe el ejercicio del bloqueo gubernamental; ciertos gobiernos imponen restricciones en sitios web y servicios en línea para controlar el acceso a la información y limitar el disenso⁸. Existen presiones económicas, donde empresas y anunciantes pueden influir en la modelación de los contenidos, lo que lleva a la eliminación de opiniones incómodas y la publicación solo de contenidos favorables. Además, las presiones gubernamentales piden a las plataformas que eliminen publicaciones que consideran peligrosas o ilegales. En ciertos países, las autoridades pueden exigir la suspensión de cuentas de activistas o periodistas. Se han llegado a proponer leyes para obligar a las plataformas a modificar sus algoritmos y evitar la viralización de ciertos temas. Como ejemplo sirva lo ocurrido durante la pandemia de COVID-19, varias plataformas digitales aplicaron censura a cuentas que difundían información considerada como desinformación o que cuestionaban las narrativas oficiales. Un ejemplo claro fue la eliminación de contenido en YouTube. También se reportaron casos de



8/ García Morales, M. J. La prohibición de la censura en la era digital. Universidad Autónoma de Barcelona. UNED. *Teoría y Realidad Constitucional*, n.º 31, 2013, pp. 237-276.

censura algorítmica, donde publicaciones sobre protestas contra restricciones sanitarias fueron ocultadas o desaceleradas en redes sociales. En algunos países, medios independientes denunciaron bloqueos de cuentas y restricciones en la difusión de contenido crítico con las medidas gubernamentales. En este contexto, la controversia en torno a la regulación de estos espacios tiene que ver con la característica central de su modelo de funcionamiento: dejar que los contenidos, informaciones y mensajes se puedan publicar y visibilizar sin controles editoriales previos (o *ex ante*) como ocurría en el modelo de los medios de comunicación tradicional⁹. Diversos análisis han mostrado que, antes de la adquisición de Twitter por Elon Musk el 27 de octubre de 2022, la plataforma ya enfrentaba presiones significativas por parte de gobiernos que buscaban limitar contenidos críticos, especialmente en países como Turquía e India, donde las autoridades amenazaron incluso con cerrar el servicio si no se eliminaba información sensible, como ocurrió durante las protestas agrícolas en India y en vísperas de elecciones en Turquía. La plataforma fue modulando sus políticas para gestionar la «objeción» y los límites de la expresión en contextos de crisis y controversia pública (De Keulenaar, Magalhães y Ganesh, 2023). Tras la compra y liberalización discursiva impulsada por Musk, caracterizada por la eliminación de políticas contra la desinformación y un giro explícito hacia la reducción de restricciones editoriales en X, algunos detectan un incremento sostenido de lo que consideran contenidos de desinformación y discurso de odio (Hickey, Fessler, Lerman y Burghardt, 2025). A la liberalización de X, Meta respondió reorientando sus propias estrategias de moderación: abandonó su programa de verificación de hechos, criticado por introducir sesgos y ejercer una censura excesiva, y adoptó un sistema de *community notes* inspirado directamente en el modelo aplicado por Musk en X, como parte de un giro más amplio hacia la «libre expresión» que rechazaba presiones institucionales para endurecer la censura. La eliminación del programa de *fact-checking* y su sustitución por un sistema de *community notes* inspirado en X, fue anunciado el 8 de enero de 2025 por Mark Zuckerberg. Esta reacción de Meta subraya cómo la liberalización de X reconfiguró el equilibrio competitivo y normativo en el ecosistema digital, al promover un modelo menos intervencionista que otras plataformas se vieron forzadas a considerar frente a expectativas políticas divergentes. En este nuevo equilibrio competitivo, el reposicionamiento de las grandes plataformas, incluida Meta, puede leerse a la luz de cómo dichas empresas articulan y reconfiguran públicamente sus principios de gobierno (expresión, comunidad, seguridad, elección y mejora), trasladando parte de la responsabilidad al usuario y reinterpretando valores según intereses estratégicos (Scharlach, Hallinan y Shifman, 2024). Estos son casos de quiebro en el recorrido de control que se estaba configurando. La presión sobre las grandes plataformas para intensificar o justificar políticas de moderación no provino únicamente de gobiernos nacionales, sino también de organismos supranacionales, especialmente en el caso europeo. Tanto Meta como X estuvieron sujetos a vigilancia reforzada por parte de instituciones internacionales, preocupadas por la desinformación, la integridad electoral y la seguridad digital. En el caso de X, tras los cambios impulsados por Elon Musk, la Comisión Europea abrió investigaciones formales por presuntas violaciones de las normas de moderación en el marco del Digital Services Act (DSA), exigiendo documentos internos, explicaciones sobre sus sistemas de recomendación y políticas aplicadas contra la desinformación. La Comisión ordenó a X conservar documentación interna y someterse a nuevas obligaciones de transparencia, evidenciando una presión supranacional clara y directa. Meta, por su parte, afrontó una presión similar



9/ Peña, P. (2020). Informe evidencia censura algorítmica de plataformas digitales y ataques online a medios disidentes. Informe 4. Universidad de Chile. <https://fcei.uchile.cl/noticias/173946/informe-evidencia-censura-digital-y-ataques-online-a-medios-disidentes>

por parte de la Unión Europea, que exigía un estricto cumplimiento del DSA y mayor control sobre contenidos ilegales o desinformativos. Aunque Meta anunció inicialmente su giro hacia modelos menos restrictivos, la Comisión Europea respondió intensificando su escrutinio sobre la empresa, lo que demuestra que los organismos supranacionales veían con preocupación la tendencia hacia la liberalización de políticas inspiradas en Musk. Google también ha estado sometido a presiones institucionales y supranacionales; en el contexto del endurecimiento regulatorio europeo, Google se vio obligado a acatar exigencias vinculadas al DSA, similares a las impuestas a Meta. Google ha debido «acceder» o «acatar» las exigencias regulatorias europeas para evitar sanciones y cumplir con las normas de protección de los usuarios. En los últimos años, la industria audiovisual infantil ha experimentado un proceso creciente de revisión, actualización y reedición digital de contenidos con el objetivo de alinearlos con nuevas sensibilidades sociales, especialmente aquellas vinculadas a la representación de género y diversidad. Este fenómeno se ha intensificado en paralelo al aumento de investigaciones y debates sobre el impacto que las narrativas mediáticas ejercen en la construcción de los roles de género durante la infancia, un aspecto ampliamente documentado por los estudios de socialización mediática y análisis de contenido audiovisual. El informe reciente del Geena Davis Institute muestra, por ejemplo, que las representaciones de género en la televisión infantil continúan reproduciendo brechas significativas en la distribución de personajes, lo que impulsa a la industria a modificar sus estándares y a revisar obras nuevas y antiguas para evitar perpetuar modelos restrictivos o estereotipados. La revisión de contenidos no solo afecta a producciones contemporáneas, sino también a títulos clásicos que, pese a su popularidad, pueden contener representaciones hoy consideradas problemáticas por asociaciones de familias, organismos reguladores o plataformas de *streaming*. Actualmente, las plataformas revisan guiones, sustituyen términos, ajustan diálogos y, en algunos casos, reeditan digitalmente voces o imágenes para actualizar códigos de representación y evitar la perpetuación de modelos considerados excluyentes. Herramientas digitales permiten cambiar voces, regrabar diálogos sin necesidad del elenco original, ajustar diseños de personajes y eliminar o reemplazar fragmentos que puedan resultar problemáticos. La existencia de tecnologías de transformación facial y corporal, como las empleadas en aplicaciones de *gender swap* impulsadas por algoritmos generativos, demuestra la capacidad técnica actual para alterar identidades, rasgos y apariencias de personajes de forma masiva y precisa. Las plataformas «actualizan» contenidos para promover valores inclusivos, mientras que sectores denuncian estos cambios como una reescritura ideológica del entretenimiento infantil. Disney retiró digitalmente del filme a *Sunflower*, *Fantasia* (1940) una *centaurette* negra representada con rasgos caricaturescos de *minstrel show*. Fue removida de las ediciones de Disney+ para eliminar esta representación racial problemática. De *Toy Story 2* (1999), en la versión para Disney+ (y ya desde el Blu-ray 2019), se retiró completamente la escena en la que Stinky Pete coquetea con dos Barbies en un *casting couch*. Fue eliminada por considerarse inapropiada para audiencias infantiles actuales. En *Santa's Workshop* (1932), el corto de animación clásico fue modificado en Disney+ para eliminar muñecas negras caricaturizadas, consideradas ofensivas, aunque dejaron otras figuras estereotípicas. Netflix ofreció una versión censurada de *Back to the Future II* (1989) donde se recortó la escena de revista *Ooh La La*, eliminando planos considerados inadecuados. Luego restauró la versión original tras la polémica que se generó. Usuarios detectaron que episodios de *Sonny With a Chance* (2009), disponibles en plataformas, habían sido recortados y ciertos segmentos desaparecieron. La eliminación selectiva de declaraciones en cadenas televisivas y plataformas digitales se ha convertido en un fenómeno creciente dentro del ecosistema mediático, especialmente en contextos de alta polarización política. En el ámbito televisivo, diversos episodios



recientes revelan cómo entrevistas o fragmentos completos son suprimidos o alterados, desencadenando acusaciones de manipulación informativa. Un caso paradigmático es el de CBS News, donde la cadena afrontó investigaciones y críticas por supuestamente editar o suprimir partes de entrevistas políticas. El 9-10 de noviembre de 2025 dimiten el director general de la BBC, Tim Davie, y la jefa de noticias Deborah Turness, tras la polémica por la edición de un discurso de Trump del 6 de enero de 2021 en un documental de la serie Panorama, emitido justo antes de las elecciones estadounidenses de 2024. Paralelamente, en los medios digitales y redes sociales, la supresión de declaraciones adquiere otra dimensión: plataformas como X o TikTok se han visto envueltas en acusaciones de retirar vídeos, testimonios o fragmentos de entrevistas por presunta «desinformación» o por vulnerar normas comunitarias, reforzando la percepción de que tanto empresas tecnológicas como cadenas tradicionales ejercen un control cada vez más intenso sobre qué expresiones, testimonios o versiones de los hechos llegan al público. La dinámica contemporánea de silenciamiento mediado por tecnologías digitales puede entenderse como una nueva *Espiral del silencio* (Noelle-Neumann, 1974), pero ahora reformulada bajo las lógicas algorítmicas y procedimentales de las plataformas digitales. En su versión actual, el silenciamiento opera además a través de mecanismos técnicos que borran imágenes, eliminan sonidos, recortan entrevistas y hacen desaparecer fragmentos de la realidad misma. En España, ejemplos como el del caso Rubiales, donde el borrado o supresión de declaraciones públicas se ha convertido en un elemento central del debate social, particularmente desde que la resolución judicial del 20 de febrero de 2025 situó en primer plano cómo distintas versiones y declaraciones del expresidente de la RFEF fueron apareciendo y desapareciendo del espacio público en medio del proceso mediático y judicial (Audiencia Nacional, 2025); este fenómeno encuentra paralelismos en la conducta de responsables políticos que eliminan rastros digitales, como la decisión del nuevo alcalde de Nueva York de borrar todas las publicaciones en X del anterior alcalde (McDonough, 2026), Eric Adams, al asumir su cargo en enero de 2026, o los casos de dirigentes que eliminan mensajes y registros telefónicos, como el escándalo del alcalde de Minneapolis, Jacob Frey, acusado de suprimir mensajes de texto relacionados con su actividad oficial en 2025 (Red Lake Nation News, 2025). Este último caso de autocensura por preservar intereses propios tiene su paralelismo en España, donde los debates recientes sobre la conservación de comunicaciones vinculadas a altos cargos, incluido el fiscal general, han girado igualmente en torno a la importancia de garantizar la integridad de la información pública, la preservación de dispositivos y la trazabilidad digital documental cuando existen investigaciones o procedimientos abiertos. Ya no se trata únicamente de aquel *verba volant*¹⁰ que remitía a la fragilidad de la palabra hablada, sino de una mutación más profunda en la que la propia realidad «vuela» o se volatiliza, suprimida o reconfigurada en tiempo real por intermediarios digitales. Este proceso puede producirse siguiendo, en ocasiones, los dictámenes de agencias supuestamente verificadoras, cuyas decisiones, aunque orientadas a la moderación o corrección informativa, acaban ejerciendo un poderoso efecto de silenciamiento estructural y condicionando qué expresiones, imágenes o narrativas permanecen accesibles (Musicco-Nombela, Fresneda y Zazo, 2024). Todos estos casos muestran de qué manera la censura digital planea sobre la comunicación y, en ocasiones, como rapaz en caída, la amordaza.

10/ La locución latina *verba volant, scripta manent* puede traducirse como «las palabras se las lleva el viento, lo escrito perdura». Se atribuye a Cayo Tito, quien la pronunció en un discurso ante el Senado romano. La frase enfatiza la naturaleza efímera de lo dicho en contraste con la durabilidad de lo registrado por escrito. Bien puede aplicarse a la naturaleza efímera digital y alterable.

3. Verdad y mentira en la era de la inteligencia artificial

La censura en los sistemas de inteligencia artificial puede ejercerse de manera estructural, incorporada directamente en el diseño de los modelos y en los datos con los que se entrenan. En esta forma de control, la censura no opera como una intervención posterior, sino como una limitación previa, ya que la selección, filtrado y exclusión de ciertos textos, imágenes o discursos durante el entrenamiento determina qué contenidos existen, y cuáles desaparecen, dentro del universo cognitivo del modelo. Esta forma de censura puede consolidarse sin que el usuario la perciba, porque el sistema nunca llega a generar respuestas que contradigan los criterios impuestos. La intervención puede proceder tanto de decisiones técnicas (garantizar la seguridad, evitar sesgos dañinos) como de criterios normativos, corporativos o regulatorios, que definen qué expresiones son permitidas y cuáles deben ser invisibilizadas. Así, el poder de decidir «qué entra» en la inteligencia artificial es un mecanismo de silenciamiento tan efectivo como el borrado explícito, pues configura el entorno epistémico desde el origen. La censura también puede ejercerse de manera operativa, es decir, mediante mecanismos que actúan en la interacción cotidiana entre el usuario y el sistema. Aquí, la inteligencia artificial aplica filtros, bloqueos, reformulaciones o desvíos temáticos cuando detecta contenidos considerados inapropiados, sensibles o políticamente problemáticos. Este tipo de censura no elimina la información del modelo, pero sí la interrumpe, modula o reconduce, de forma que ciertos discursos nunca llegan plenamente al usuario. A ello se suman los sistemas de moderación externa, como agregadores, plataformas y verificadores, que pueden establecer restricciones adicionales y determinar qué respuestas son visibles, accesibles o amplificadas. En conjunto, estas capas de control configuran una nueva forma de censura dinámica y distribuida, donde múltiples actores, humanos y algorítmicos, intervienen en tiempo real para delimitar el perímetro expresivo de la inteligencia artificial.

En la era de la inteligencia artificial, la censura digital adquiere una nueva densidad epistemológica al intervenir directamente en los procesos mediante los cuales se construyen la verdad y la mentira en el espacio público. Los sistemas de inteligencia artificial no solo filtran información, sino que determinan qué versiones de la realidad permanecen visibles y cuáles se desvanecen, modulando así la frontera entre lo verificable y lo negado. Esta censura ya no actúa únicamente suprimiendo contenidos, sino reordenando jerarquías de credibilidad, privilegiando narrativas aprobadas por algoritmos o entidades verificadoras y relegando otras a la invisibilidad técnica. En este nuevo régimen informacional, la mentira puede naturalizarse si los sistemas que median el acceso al conocimiento la reproducen o no la corrigen, mientras que la verdad puede volverse irreconocible si el aparato digital que la sostiene la silencia, la oculta o la etiqueta como irrelevante. De este modo, la inteligencia artificial convierte la noción clásica de verdad, como correspondencia con los hechos, en una cuestión de accesibilidad y persistencia digital, y transforma la mentira, antes un acto de voluntad, en un fenómeno de distribución algorítmica, donde lo que existe es aquello que las plataformas permiten circular. Así, por ejemplo, en esta nueva ecología comunicativa donde la verdad se confunde con la mera persistencia mediática, incluso los dramáticos hechos del accidente ferroviario ocurrido en España en enero de 2026 pueden verse neutralizados: cuando los responsables políticos comparecen repetidamente reconociendo «errores», la insistencia performativa de su presencia, aunque hayan mentido o gestionado de forma engañosa, acaba funcionando como un sello de transparencia, de modo que su mentira queda absorbida por la apariencia de verdad que produce el acto mismo de comparecer, sin que existan ya consecuencias reales sobre la falsedad aplicada a los hechos.



4. Resultados

Los resultados permiten identificar un conjunto de mecanismos de censura digital que operan de forma estructural, operativa y forense, reconfigurando la visibilidad y el estatuto público de la verdad en el ecosistema comunicativo contemporáneo. La tabla 1 sintetiza estas tipologías, sus ejemplos más significativos y las implicaciones epistémicas y éticas derivadas de cada una.

Nuestro análisis muestra que la censura digital contemporánea funciona menos como tijera explícita y más como arquitectura de visibilidad: decide qué aparece, cómo aparece y qué desaparece del campo perceptivo público. En esa línea, Vicente García Escrivá (2026) describe el

Tabla 1. Tipologías de censura digital: mecanismo, ejemplos, daño a la verdad y consecuencias éticas

Tipo de censura digital	Mecanismo/ Descripción	Ejemplos concretos	Daño a la verdad (epistémico)	Consecuencias éticas/profesionales
Algorítmica estructural	Selección/ expulsión de datos en entrenamiento y reglas de recomendación que impiden que ciertos contenidos existan o circulen	Bloqueo de temas «sensibles» por <i>chatbots</i> ; reglas de ranking que invisibilizan fuentes o perspectivas	La verdad pierde existencia pública al no ser accesible ni generable; se confunde lo verdadero con lo persistente/ recomendado	Falta de transparencia y rendición de cuentas ; asimetría de poder técnico; normalización del silenciamiento
Operativa (filtros/ moderación)	Bloqueos, reformulaciones o desvíos temáticos en la interacción; etiquetado como «desinformación/ odio»	Filtros de plataforma y colaboraciones con verificadores; limitación de reenvíos o difusión	Reordena jerarquías de credibilidad : lo etiquetado pierde verosimilitud; lo no etiquetado gana aura de verdad	Riesgo de arbitrariedad ; opacidad procedimental; refuerzo de la espiral del silencio
Shadow banning/ despriorización	La pieza existe, pero no aparece en búsquedas/ <i>feeds</i> ; alcance reducido por <i>ranking</i>	Desaceleración de protestas/temas «conflictivos»; menor exposición de determinadas cuentas	La verdad se vuelve inaccesible : sin circulación no hay contraste público; relevancia ≠ verdad, sino posición en ranking	Vulneración del pluralismo ; incentiva autocensura ; deterioro del debate
Bloqueo/ suspensión de cuentas	Cierres automáticos o por revisión manual por supuestas infracciones	Suspensiones por «odio»/ «desinformación»; bloqueos a periodistas/ activistas en algunos países	Pérdida de testimonios clave y de memoria documental ; empobrecimiento del contraste de fuentes	Erosión de la libertad de expresión ; efecto desaliento (chilling effect)
Geo filtrado y localización forzada	Resultados restringidos por IP/ idioma; sustitución de ccTLD por dominio global con geolocalización	Redirección a resultados locales; menor acceso a contenidos transnacionales e idiomas alternativos	Encierro informativo y provincialización del horizonte de verificación; menor comparabilidad de fuentes	Limita autonomía

Fuente: Elaboración propia.



tránsito desde los códigos restrictivos del *noir* clásico hacia la lógica de las plataformas, donde el «deseo de ver» queda mediado por algoritmos que operan como nuevos clasificadores: no solo filtran, sino que curan el encuadre de lo visible, convierten la moderación en un régimen de apariencia y administran la escasez de atención. Esta curaduría algorítmica, lejos de ser neutra, impone jerarquías de credibilidad y de relevancia que reordenan silenciosamente el mapa de lo que cuenta y lo que queda fuera. En paralelo, Jacinto Gómez y Gema Pérez (2026) documentan que el borrado de publicaciones, la autocensura preventiva y el lavado de imagen constituyen una economía del «error controlado»: se editan y suprimen rastros, se repiten comparencias y se performa la disculpa hasta que la insistencia de la presencia se confunde con transparencia, desplazando la carga de la prueba desde los hechos hacia la escenografía del reconocimiento. El resultado es una paradoja: el acto de decir termina erigiéndose en evidencia de verdad allí donde los contenidos (o su gestión) pudieron ser incompletos o engañosos. A partir de ahí, como propone José Díaz Cuesta (2026), conviene entender la censura actual como un fenómeno multicapa: estructural (qué datos entran a los modelos y qué reglas de recomendación se imponen), operativa (qué filtros de generación bloquean o reformulan respuestas en tiempo real) y forense (qué reescrituras *ex post* eliminan, etiquetan o despriorizan piezas ya publicadas). Estas capas producen una metamorfosis: la verdad pública no desaparece necesariamente por prohibición directa, sino porque pierde circulación, queda degradada en el *ranking* o es rodeada de señales de desconfianza que desincentivan su lectura, mientras que ciertos enunciados falaces ganan verosimilitud por saturación repetitiva y por la apariencia de legitimidad que les confieren los mismos mecanismos de distribución. La brecha ya no es solo entre lo verdadero y lo falso, sino entre lo accesible y lo inaccesible, lo amplificado y lo silenciado. En este desplazamiento epistemológico, la clásica definición de verdad como correspondencia con los hechos se ve tensionada por su persistencia algorítmica: lo que no circula, lo que no se recomienda, lo que se borra o aparece descontextualizado, deja de existir para efectos cívicos. Por eso, el enfoque humanista de Encarnación Fernández Gómez (2026), que concibe el conocimiento como viaje, se vuelve crucial: ver no equivale a tener acceso técnico, sino a aprender a mirar, a experimentar el encuadre, reconocer el montaje, detectar los silencios y reconstruir el contexto. En ese mismo espíritu, el análisis lingüístico de Iryna Moysenko (2026) sobre tácticas de manipulación (ambigüedad estratégica, desplazamientos semánticos, apelaciones emocionales) explica por qué, en un ecosistema algorítmico, la mentira ya no opera solo como acto voluntario, sino como fenómeno de distribución: lo que se comparte, se refuerza y se replica adquiere aura de verosimilitud independientemente de su correspondencia con los hechos. Los datos sugieren un nuevo régimen de censura digital que gestiona la visibilidad (qué entra y con qué marco), administra el error (borrados selectivos y disculpas performativas) y reingeniería la credibilidad (verdad y mentira reordenadas por reglas de prioridad y acceso). De ahí se desprende una implicación decisiva para nuestro campo: la educación no puede limitarse a la alfabetización informacional básica, sino que ha de ser experiencial y crítica. La experiencia del proyecto de innovación docente que presentas constituye, en este sentido, un laboratorio formativo: al poner a los estudiantes frente a casos reales, a prácticas de edición, a experimentos con *prompts* y a auditorías de trazabilidad, aprenden a ver el encuadre, a argumentar con evidencia y a documentar los procesos que hacen posible, o imposible, la verdad pública. Esa práctica guiada, sostenida y reflexiva no solo refuerza competencias analíticas; desactiva la nueva espiral del silencio digital, restituye el valor cívico de la verdad y compromete a las instituciones (y a las plataformas) con responsabilidades concretas de custodia, contexto y transparencia.

A partir de estos hallazgos, la educación se revela como el antídoto imprescindible: no solo una alfabetización mediática general, sino una didáctica experiencial que ponga al alum-



nado en situación de detectar, analizar, auditar y producir; que les permita probar cómo cambia el conocimiento sobre los hechos y el acercamiento a la verdad a través del desarrollo del pensamiento crítico. En este sentido, la experiencia de proyectos de innovación docente llevados a cabo en universidades como *Ni corruptores, ni corruptos, ni maquinistas del fango: buscar y comunicar la verdad* (NCNM-BV, 2025)¹¹, por su enfoque aplicado, el trabajo con casos reales y la integración de análisis técnico-lingüístico con reflexión ética, constituye un laboratorio de aprendizaje donde el estudiantado ensaya competencias críticas de alto nivel: saber ver, saber argumentar y saber documentar.

La constatación contemporánea de que la censura digital ya no opera mediante prohibiciones explícitas, sino a través de arquitecturas invisibles de visibilidad, borrado o amplificación algorítmica, encuentra un eco profundo en el proyecto interdisciplinar *Ni corruptores, ni corruptos, ni maquinistas del fango*. Si, como hemos visto, los algoritmos modulan qué hechos permanecen accesibles y cuáles se diluyen en un ruido hiperemocional, si las comparencias políticas reiteradas transforman la mentira gestionada en apariencia de transparencia, y si la reiteración comunicativa sustituye a la veracidad como nuevo criterio de credibilidad, entonces nuestro ámbito educativo se convierte en un espacio crucial para reconstruir las condiciones mismas del juicio ético y profesional. Este proyecto asume precisamente ese desafío: formar estudiantes capaces de reconocer la diferencia entre lo verdadero y lo verosímil, entre la gestión estratégica del error y la responsabilidad moral, entre la edición interesada de los hechos y la honestidad narrativa. Frente a un ecosistema donde la verdad se hace frágil y la mentira se vuelve eficiente, la experiencia docente propuesta articula una pedagogía capaz de poner al alumnado en contacto directo con aquellos mecanismos de manipulación, simulación y reconstrucción del relato que dominan hoy la comunicación política, empresarial y periodística. La creación de mundos ficcionales inspirados en dinámicas reales, campañas políticas, crisis reputacionales, decisiones jurídicas, tensiones mediáticas, permite que cada estudiante experimente en primera persona aquello que de otra forma solo percibiría como espectador distante: el peso del encuadre, el efecto de la edición, la tentación de la autocensura, la responsabilidad de sostener un relato veraz incluso bajo presión, la dificultad de identificar tácticas de manipulación lingüística o emocional y el riesgo de convertirse, sin advertirlo, en un agente reproductor del fango comunicativo. La combinación de simulación, narrativa transmedia, *role-play*, pedagogía de las virtudes y análisis crítico interdisciplinar convierte al aula en un laboratorio donde la verdad vuelve a ser un problema vivo, a veces incómodo, siempre exigente. En este sentido, el proyecto responde directamente a la crisis descrita: allí donde la censura algorítmica borra, el proyecto enseña a reconstruir; donde la sobreexposición convierte la mentira en acto performativo, el proyecto enseña a discernir; donde el cinismo profesional normaliza el relato por encima de los hechos, el proyecto enseña a servir a la verdad como virtud ética y como responsabilidad pública. Así, esta iniciativa no solo dota al alumnado de competencias técnicas en comunicación política, empresarial y periodística, sino que lo forma como una comunidad de profesionales capaces de resistir la deriva manipulativa de nuestro tiempo y de restaurar un paradigma comunicativo valiente, ético y orientado al bien común. Solo desde esa práctica guiada, sostenida y reflexiva es posible desactivar la nueva espiral del silencio digital, restituir el valor público de la verdad y responsabilizar a instituciones y plataformas de sus decisiones de visibilidad.

11/ Universidad Francisco de Vitoria (2025). *Ni corruptores, ni corruptos, ni maquinistas del fango: buscar y comunicar la verdad* (NCNM-BV). Proyecto seleccionado en la Convocatoria de Ayudas a la Innovación en el Aprendizaje 2025-2026. <https://innovacionufv.com/convocatoria-de-ayuda-a-la-innovacion-en-el-aprendizaje-2025-2026/>

5. Conclusiones

Al final, solo la justicia y la ética de quienes la defienden y la aplican pueden garantizar la búsqueda de la verdad y el derecho a la libertad de comunicación, estableciendo límites a la censura. Como establece el principio jurídico: «Solo podrá acordarse el secuestro de publicaciones, grabaciones y otros medios de información en virtud de resolución judicial». La censura es un acto humano y, por ello, el poder de restringirla o perpetuarla reside en quienes la ejercen.

Para concluir, es necesario hacer una última llamada de atención en esta era digital, marcada por el auge de la inteligencia artificial y la progresiva desaparición de lo material. Nos enfrentamos a la llamada modernidad líquida, donde lo digital es vulnerable a la manipulación y al borrado absoluto. En este nuevo paradigma, nada queda registrado de forma definitiva. Hoy, con una facilidad inquietante, textos, documentos y declaraciones pueden desaparecer sin dejar rastro. La era digital permite borrar y alterar la información como nunca antes en la historia de la humanidad, desde los antiguos papiros hasta nuestros días. Nos enfrentamos a una nueva forma de censura que opera sin restricciones, eliminando o modificando lo incómodo mediante sofisticadas herramientas *deepfake*¹². Cuidado con las supuestas democracias digitales y los sufragios universales a través de aplicaciones móviles; el poder de alterar la verdad nunca ha sido tan accesible.

La restauración de la verdad pública en la era digital exige una vuelta a la interioridad crítica que ordene la mirada, el juicio y la acción. Como recordaba san Agustín, «No vayas fuera; vuelve a ti mismo. En el hombre interior habita la verdad; y si descubres que tu naturaleza es mutable, trasciéndete a ti mismo» (*Noli foras ire, in teipsum redi; in interiore homine habitat veritas ... transcende et teipsum*)¹³. Esta llamada a reentrar para discernir, antes de hablar, distribuir o amplificar, es el principio ético que debe coronar cualquier alfabetización mediática, cualquier innovación docente y toda práctica profesional orientada al bien común.



Bibliografía

- Agustín de Hipona (s. f.). De vera religione, 39, 72-73. En Augustinus Hipponensis - De Vera Religione liber unus. https://www.augustinus.it/latino/vera_religione/vera_religione.htm
- Agustín de Hipona (2007). Confesiones (ed. y trad. esp.). Madrid: Biblioteca de Autores Cristianos. (Obra original 397-400).
- Audiencia Nacional, Juzgado Central de lo Penal. (2025, 20 de febrero). Sentencia n.º 3 (Proc. Abreviado 6/2024; Diligencias Previas 62/2023): Luis Manuel Rubiales Béjar y otros. https://s1.elespanol.com/2025/02/20/actualidad/2025-02-20_SENTENCIA_RUBIALES.pdf
- Código de buenas prácticas reforzado sobre desinformación (2022). Comisión Europea. Strengthened Code of Practice on Disinformation. <https://digital-strategy.ec.europa.eu/en/policies/code-practice-disinformation> y texto completo: https://disinfocode.eu/assets/pdfs/2022_Strengthened_Code_of_Practice_Disinformation.pdf

12/ Tecnologías, conocidas como *deepfake* de voz, utilizan inteligencia artificial y aprendizaje profundo para analizar patrones vocales y reproducirlos con alta fidelidad; permiten imitar voces humanas con gran precisión, generando audios que pueden hacer que una persona parezca decir algo que nunca ha dicho.

13/ Agustín de Hipona (s. f.). De vera religione, 39, 72-73. En Augustinus Hipponensis - De Vera Religione liber unus (texto latino). Recuperado de https://www.augustinus.it/latino/vera_religione/vera_religione.htm

- Constitución Española (1978). Boletín Oficial del Estado, n.º 311, 29 de diciembre de 1978. <https://www.boe.es/buscar/pdf/1978/BOE-A-1978-40001-consolidado.pdf>
- De Keulenaar, E., Magalhães, J. C., y Ganesh, B. (2023). Modulating moderation: A history of objectionability in Twitter moderation practices. *Journal of Communication*, 73(3), 273-287. <https://doi.org/10.1093/joc/jqad015>
- Díaz Cuesta, J. (2026). La metamorfosis de la censura en la era digital: Desafíos algorítmicos. *Comunicación y Hombre*, (22), 9-15. <https://doi.org/10.9999/comyh.2026.01>
- DSA. Digital Services Act (2022). Ley de Servicios Digitales Reglamento (UE) 2022/2065, del Parlamento Europeo y del Consejo, de 19 de octubre de 2022. Diario Oficial de la Unión Europea, L 277, 1-102. <https://eur-lex.europa.eu/eli/reg/2022/2065/oj>
- Fernández Gómez, E. (2026). Caleido. El mito de la torre bella: viaje hacia el conocimiento. *Comunicación y Hombre*, (22), 16-22. <https://doi.org/10.9999/comyh.2026.02>
- García Escrivá, V. (2026). Regímenes de lo visible: censura y deseo de ver del noir clásico a la era del algoritmo. *Comunicación y Hombre*, (22), 23-29. <https://doi.org/10.9999/comyh.2026.03>
- Geena Davis Institute on Gender in Media. (2004-). Geena Davis Institute on Gender in Media. <https://geenadavisinstitute.org/>
- Hickey, D., Fessler, D. M. T., Lerman, K., y Burghardt, K. (2025). X under Musk's leadership: Substantial hate and no reduction in inauthentic activity. *PLOS ONE*, 20(2), e0313293. <https://doi.org/10.1371/journal.pone.0313293>
- Konstantinovic, D. (2025, enero 8). Meta will ease up on content moderation in significant reversal. eMarketer. <https://www.emarketer.com/content/meta-will-ease-up-content-moderation-significant-reversal>
- Masini, T. (2015). Testa bendata [Obra pictórica]. Artmajeur.
- McDonough, A. (2026, January 2). Mamdani plans to delete ALL Adams' tweets. City & State New York. <https://www.cityandstateny.com/policy/2026/01/mamdani-plans-delete-all-adams-tweets/410435/>
- Moyseyenko, I. (2026). Análisis lingüístico de las noticias falsas: Detección de tácticas de manipulación. *Comunicación y Hombre*, (22), 37-43. <https://doi.org/10.9999/comyh.2026.05>
- Musicco-Nombela, D., Martínez-Fresneda, H., y Zazo Correa, L. (2024). La universidad como lugar para fortalecer la libertad en la comunicación: Combatir la nueva censura en la era digital y en las democracias, *Pensar sobre la comunicación: Universidad y empresa* (pp. 241-258). Tirant lo Blanch.
- Noelle-Neumann, E. (2010). La espiral del silencio: Opinión pública: nuestra piel social (F. J. Ruiz Calderón, Trad.). Barcelona: Paidós.
- Pérez, G. y Gómez, J. (2026). Borrado de post, autocensura y lavado de imagen. *Comunicación y Hombre*, (22), 30-36. <https://doi.org/10.9999/comyh.2026.04>
- Quartz (2024, septiembre 25). Elon Musk criticizes censorship, but his X grants more government requests than Twitter did. <https://qz.com/elon-musk-x-twitter-transparency-brazil-turkey-tesla-1851656960>
- Red Lake Nation News (2025, October 16). Political group accuses Frey of improperly deleting texts. Some advocates and media lawyers agree. <https://www.redlakenationnews.com/story/2025/10/16/politics/political-group-accuses-frey-of-improperly-deleting-texts-some-advocates-and-media-lawyers-agree/135297.html>
- Scharlach, R., Hallinan, B., y Shifman, L. (2024). Governing principles: Articulating values in social media platform policies. *New Media & Society*, 26(11), 6658-6677. <https://doi.org/10.1177/14614448231156580>
- Universidad Francisco de Vitoria (2025). Ni corruptores, ni corruptos, ni maquinistas del fango: buscar y comunicar la verdad (NCNM-BV). Proyecto seleccionado en la Convocatoria de Ayudas a la Innovación en el Aprendizaje 2025-2026. <https://innovacionufv.com/convocatoria-de-ayuda-a-la-innovacion-en-el-aprendizaje-2025-2026/>

