VISUAL REVIEW | Vol. 18, No. 4, 2026 | ISSN 2695-9631
International Visual Culture Review / Revista Internacional de Cultura Visual
https://doi.org/10.62161/revvisual.v18.6396

# ESTÉTICA DE LA DESINFORMACIÓN EN LA CIRCULACIÓN NOTICIOSA DIGITAL
## Análisis multimodal de pruebas visuales falsas

HENRI EMMANUEL LOPEZ GOMEZ [1]
C31648@utp.edu.pe
[1]Universidad Tecnológica del Perú, Perú

**PALABRAS CLAVE**

*Desinformación visual*
*Análisis multimodal*
*Evidencialidad*
*Recontextualización*
*Deepfakes*
*Verificación*
*Montajes compositivos*

**RESUMEN**

*Este estudio analiza la estética de la desinformación en noticias a partir de 60 piezas visuales falsas (capturas recontextualizadas, montajes compositivos y contenidos sintéticos) recolectadas en el ecosistema digital en español (2024-2025). Mediante análisis multimodal, se identificaron cinco conjuntos de indicios de evidencialidad (procedencia, urgencia, interfaz, señalamiento, documentalidad) y patrones de anclaje texto-imagen que construyen la apariencia de prueba. La "estética de prueba" opera como ensamblaje de marcas de procedencia, diseño noticioso y guías de lectura, ofreciendo criterios para la verificación y alfabetización visual.*

Recibido: 02/ 04 / 2026
Aceptado: 30/ 06 / 2026

## 1. Introducción

La investigación sobre desinformación en ecosistemas digitales ha crecido de forma sostenida y ha consolidado un campo interdisciplinar que conecta comunicación, ciencias cognitivas, estudios de plataformas y ciencia de datos (Broda & Strömbäck, 2024). En este marco, misinformation se define como información falsa o engañosa difundida sin intención necesariamente dañina, mientras que disinformation se asocia a la producción o circulación estratégica de falsedades con fines instrumentales (Zeng & Brennen, 2023).

La desinformación visual se entiende como un subtipo de desorden informativo en el que imágenes o vídeos —alterados, sintetizados o recontextualizados— funcionan como "evidencia" persuasiva para sostener afirmaciones falsas o equívocas (Heley et al., 2022). Dado que estas piezas combinan recursos icónicos, lingüísticos y contextuales, el análisis multimodal ha permitido describir cómo se ensamblan significados y cómo operan los indicios de autenticidad en distintos entornos de circulación (Wilson et al., 2023).

En el periodismo y la comunicación pública, las imágenes han operado como recursos de indexicalidad y prueba, pero también han facilitado distorsiones cuando se integran a narrativas engañosas en escenarios de alta incertidumbre informativa (Brennen et al., 2021).

En años recientes, la literatura ha sintetizado que la desinformación visual abarca desde recontextualizaciones "baratas" (capturas parciales, recortes, cambios de pie de foto) hasta manipulaciones complejas y contenidos sintéticos generados con IA (Weikmann & Lecheler, 2023). En paralelo, la investigación ha planteado la necesidad de una agenda específica para comprender cómo se forman los juicios de credibilidad ante "pruebas visuales" falsas y qué claves perceptivas o inferenciales se activan en la evaluación (Peng et al., 2023).

Desde la psicología del juicio, se ha establecido que las imágenes pueden modificar percepciones de verdad al aumentar la fluidez de procesamiento y la sensación subjetiva de evidencia, incluso cuando el contenido visual es no probatorio (Newman & Schwarz, 2024). En un experimento con retroalimentación de veracidad, Pulm et al. (2024) encontraron que acompañar el feedback con fotografías no probatorias mejoró la discriminación entre verdadero y falso, aunque también sesgó la tendencia a responder "verdadero" en ciertos contextos de evaluación.

En redes sociales, Hameleers et al. (2020) mostraron que la desinformación multimodal puede intensificar efectos persuasivos frente a mensajes exclusivamente textuales, en parte por su capacidad de activar respuestas afectivas y heurísticos de credibilidad.

En el caso de los deepfakes, Vaccari y Chadwick (2020) observaron que la exposición a vídeos políticos sintéticos incrementó la incertidumbre y puede erosionar la confianza, no solo en contenidos concretos, sino también en el ecosistema informativo en general.

En el plano institucional, Aljalabneh (2024) advirtió que la expansión de contenidos sintéticos tensiona las rutinas de verificación y los contratos de credibilidad entre medios, fuentes y audiencias, con efectos potenciales sobre la legitimidad del periodismo.

A nivel técnico, Mirsky y Lee (2022) documentaron una carrera armamentística entre la generación y la detección de deepfakes, en la que mejoras en realismo no implican mejoras equivalentes en la capacidad humana de identificación. Además, Rauchfleisch et al. (2025) mostraron que el encuadre terminológico (deepfakes vs. "medios sintéticos") influye en percepciones de amenaza reputacional y en actitudes hacia la regulación, lo que sugiere que la "estética" del fenómeno también se disputa discursivamente.

Las intervenciones informativas sobre autenticidad tampoco son neutras: Hameleers y Marquart (2023) hallaron que las etiquetas de deepfake pueden delegitimar discursos políticos auténticos, creando un efecto colateral de sospecha que favorece estrategias de negación plausible.

Cuando la manipulación es de baja fidelidad, Hameleers (2024) comparó cheapfakes y deepfakes y evidenció diferencias en mecanismos emocionales y cognitivos, subrayando que la alfabetización mediática y el pensamiento analítico median la susceptibilidad a engaños visuales de distinta complejidad. De forma complementaria, Tonglet et al. (2025) subrayaron que las imágenes fuera de contexto constituyen una de las formas más prevalentes de desinformación visual y requieren, para su refutación, recuperar contexto verdadero y evaluar la veracidad del pie o la afirmación asociada.

En el terreno de la gobernanza informativa, Wittenberg et al. (2025) evaluaron el impacto del etiquetado de medios generados por IA y mostraron que sus efectos sobre credibilidad y compartición dependen del diseño del aviso y del ecosistema de exposición, sin garantizar reducciones uniformes del

daño informativo. Aun así, las advertencias pueden producir un "derrame" de escepticismo: Ternovski et al. (2022) mostraron que informar sobre la existencia de deepfakes puede disminuir la confianza en noticias reales y aumentar la duda generalizada, lo que plantea dilemas para políticas de comunicación preventiva.

En el frente automatizado, Shen et al. (2024) propusieron un enfoque de detección de desinformación fuera de contexto que combina verificación encadenada y señales multimodales, reflejando un desplazamiento hacia arquitecturas que intentan "razonar" sobre coherencia entre imagen, texto y contexto. En el nivel formativo, Qian et al. (2022) demostraron que intervenciones de alfabetización digital pueden motivar el uso de búsqueda inversa y prácticas de verificación, reduciendo la vulnerabilidad ante piezas visuales recontextualizadas que simulan evidencia.

En el plano profesional, Westlund et al. (2024) conceptualizó el fact-checking como proceso de sensemaking y resolución de problemas bajo condiciones de ambigüedad, donde la verificación visual se vuelve un trabajo situado, con restricciones de tiempo, herramientas y coordinaciones interorganizacionales. En este sentido, Weikmann y Lecheler (2024) documentaron cómo periodistas y verificadores ajustan rutinas ante deepfakes entre expectativas tecnológicas, criterios editoriales y límites prácticos de autentificación, mostrando fricciones entre "soluciones" técnicas y legitimidad pública.

Asimismo, Vinhas y Bastos (2022) argumentaron que el fact-checking no solo corrige afirmaciones, sino que sostiene una "realidad de consenso" frágil, lo que vuelve especialmente relevante estudiar cómo se estabilizan —o se erosionan— las pruebas visuales en controversias públicas.

En trabajos recientes, Vinhas y Bastos (2025) analizaron la gobernanza "WEIRD" de la verificación y la moderación, subrayando que las políticas de plataforma y los regímenes de visibilidad condicionan qué se considera evidencia legítima y qué se sanciona como manipulación.

En educación, Aljalabneh (2024) identificó estrategias pedagógicas para fortalecer evaluación crítica de imágenes y vídeos —búsqueda inversa, triangulación de fuentes, análisis de edición— destacando que la potencia mnemónica de lo visual puede convertirse en vulnerabilidad si se confunde estética con prueba.

En este artículo, "contenido noticioso" se entiende de manera operativa como piezas visuales que circularon en entornos digitales presentándose explícitamente como noticia, alerta o evidencia ("última hora", "urgente", "confirmado"), o que imitaron convenciones periodísticas reconocibles (titulares, cintillos, placas informativas, logos, tipografías y marcas de interfaz). Por ello, el análisis no se restringió a productos publicados por redacciones profesionales, sino a artefactos multimodales que reclamaron estatus noticioso y activaron una lectura de "prueba" en la circulación pública.

En conjunto, la evidencia disponible ha clarificado efectos, mediadores y desafíos de la desinformación visual, pero ha tendido a tratar la imagen como "soporte" de falsedad más que como un régimen estético de prueba que imita formatos periodísticos (capturas, plantillas, rótulos, recortes, marcas de interfaz) para producir verosimilitud. En particular, permanece subanalizada la dimensión estética mediante la cual estas piezas se presentan como pruebas visuales —no solo como ilustraciones— y cómo esa estética articula autoridad, urgencia y "documentalidad" en noticias y conversaciones públicas.

El objetivo de este artículo es caracterizar la estética de la desinformación en noticias a través de las pruebas visuales falsas, proponiendo (1) una delimitación conceptual del fenómeno como estrategia de evidencialidad visual, (2) una tipología analítica de sus principales formas (capturas recontextualizadas, montajes compositivos y contenidos sintéticos), y (3) un marco de lectura multimodal para identificar los indicios de autenticidad simulada y sus implicaciones para credibilidad, verificación y cultura visual contemporánea.

## 2. Marco teórico

### *2.1. Desinformación visual y "pruebas" en el ecosistema noticioso*

La desinformación se define como información falsa o engañosa difundida con intencionalidad (p. ej., lucro, influencia política o daño), mientras que la desinformación visual se caracteriza por integrar imágenes o videos como vector central de la falsedad o el engaño, ya sea por fabricación (p. ej., contenido sintético), manipulación (montajes/edición) o descontextualización (capturas o imágenes reales

reubicadas en otro marco interpretativo). Esta distinción resulta clave porque el componente visual no funciona solo como "acompañamiento" del texto: opera como dispositivo de prueba, es decir, como soporte que pretende cerrar la brecha entre afirmación y evidencia mediante una apariencia documental (Hameleers, 2025; Hausken, 2024; Weikmann & Lecheler, 2023).

En la literatura reciente, el estudio de la desinformación visual se ha consolidado como un subcampo con problemas propios: (a) su alta circulación en entornos de plataforma; (b) su capacidad para anclar narrativas en señales perceptivas de autenticidad; y (c) su plasticidad para imitar géneros periodísticos (capturas de pantalla, "evidencias" forenses, placas informativas, clips "de archivo"). En particular, las síntesis de investigación han propuesto tratar la desinformación visual como un fenómeno multimodal, donde la credibilidad emerge de la interacción entre contenido visual, texto, contexto de publicación e indicadores de procedencia (Hameleers, 2025; Weikmann & Lecheler, 2023).

Desde este encuadre, las "pruebas visuales falsas" (capturas, montajes, deepfakes) no se reducen a un problema tecnológico; se entienden como estrategias sociosemióticas que movilizan expectativas culturales sobre qué cuenta como evidencia en noticias. Dichas expectativas se han formado históricamente en torno al estatuto testimonial de la fotografía y el video, pero hoy se ven reconfiguradas por la disponibilidad de edición avanzada, generación sintética y circulación acelerada en plataformas (Brennen et al., 2021; Hameleers, 2025; Newman & Schwarz, 2024).

## *2.2. Evidencia visual, verosimilitud y crisis de la confianza perceptiva*

En comunicación pública, la imagen suele operar como un atajo epistémico: se asume que "ver" aproxima a "saber". La investigación en psicología cognitiva ha mostrado que las imágenes influyen en juicios de verdad, muchas veces fuera de la conciencia metacognitiva, y que su efecto persuasivo no depende únicamente de que la imagen sea explícitamente manipulada: también puede darse por selección, encuadre o recontextualización. En consecuencia, el riesgo no se limita al *deepfake*; incluye un continuo de "evidencias" visuales que sostienen inferencias no justificadas (Brennen et al., 2021; Newman & Schwarz, 2024).

Un punto teórico decisivo es que el entorno contemporáneo ha erosionado el supuesto de que la apariencia fotográfica equivale a registro. El realismo visual puede ser producido por medios no indexicales (p. ej., generación por IA), lo cual desplaza la confianza desde el "origen" del registro hacia la "calidad" de la simulación. En términos estéticos, esto implica que la credibilidad se negocia mediante convenciones de fotorealismo y de género (cómo "debe verse" una evidencia noticiosa), y no solo mediante garantías de procedencia (Hausken, 2024; Newman & Schwarz, 2024).

En paralelo, la discusión filosófica reciente ha cuestionado lecturas apocalípticas y ha propuesto conceptualizar el impacto de los deepfakes como un problema de prácticas epistémicas: qué rutinas de verificación, qué umbrales de sospecha y qué economías de atención hacen que ciertos públicos acepten o rechacen evidencia audiovisual. Así, el daño central no sería únicamente la persuasión de un contenido falso, sino la inestabilidad de los criterios de confianza (p. ej., pasar de "ver para creer" a "nada es creíble"), con consecuencias para la autoridad informativa del periodismo (Habgood-Coote, 2023; Hameleers, 2025).

Finalmente, la vulnerabilidad no solo es técnica; también es humana. Evidencia experimental ha indicado que las personas tienden a sobrestimar su capacidad para detectar deepfakes y a sesgar sus juicios hacia la autenticidad ("seeing-is-believing"), lo que vuelve más eficiente la inserción de "pruebas visuales" en narrativas engañosas. Este hallazgo sustenta la necesidad de analizar la desinformación visual como un fenómeno donde percepción, cultura visual y contexto mediático interactúan (Köbis et al., 2021).

## *2.3 Semiótica social y multimodalidad como lente para analizar "pruebas" falsas*

Para estudiar cómo las "pruebas visuales" producen credibilidad, este trabajo se apoya en la semiótica social y el enfoque multimodal, que entienden la comunicación como articulación de recursos (imagen, tipografía, color, composición, lenguaje, interfaz) y consideran que el sentido no reside en un modo aislado, sino en su orquestación. En esta tradición, las formas visuales se analizan como gramáticas culturales que codifican representaciones, relaciones (p. ej., distancia, autoridad) y organización textual (jerarquías, saliencias) (G. Kress & Van Leeuwen, 2020).

Aplicado a noticias, este lente permite conceptualizar la estética de la desinformación como un mimetismo de códigos informativos: rótulos, lower thirds, capturas "con fecha", supuestos metadatos, marcas de agua, sellos de verificación apócrifos o encuadres "en vivo". No se trata solo de "mentir con imágenes", sino de construir verosimilitud mediante convenciones del periodismo audiovisual y digital. Dichas convenciones son interpretables como repertorios multimodales que los productores de desinformación reensamblan para fabricar una apariencia de evidencia (Hameleers, 2025; Weikmann & Lecheler, 2023).

Desde esta perspectiva, las categorías empíricas del estudio (capturas, montajes, deepfakes) pueden leerse como familias de procedimientos semióticos:

- Capturas recontextualizadas: desplazan el anclaje temporal y situacional, pero conservan marcas de "registro" (interfaces, timestamps, apariencia de fuente).
- Montajes: combinan indexicalidad parcial (fragmentos reales) con operaciones de edición que reconfiguran causalidad y atribución.
- Deepfakes: producen un realismo performativo (rostro/voz/movimiento) que activa expectativas de autenticidad audiovisual, aunque el origen sea sintético.

En los tres casos, la "prueba" se sostiene por la convergencia entre recursos visuales y marcos narrativos; por ello, la multimodalidad funciona aquí como un lente integrador (no como suma de variables).

### *2.4 Autenticación en plataformas: señales, atajos y disputas por la credibilidad*

La credibilidad contemporánea se produce en un entorno de plataforma, donde la evaluación de autenticidad se realiza bajo escasez de tiempo y sobreabundancia informativa. En ese contexto, las audiencias recurren a señales (cues) y atajos interpretativos: reputación de la fuente, consistencia del formato, estilo visual "profesional", presencia de enlaces, interacción social, o compatibilidad con expectativas previas. Investigaciones recientes han descrito cómo estos atajos estructuran el juicio de autenticidad en noticias online y, a la vez, abren oportunidades para la ingeniería de credibilidad por parte de actores desinformantes (Gehrke et al., 2025).

En usuarios jóvenes, por ejemplo, se ha observado que "lo auténtico" puede inferirse desde indicadores como coherencia narrativa, plausibilidad contextual y señales de legitimidad social (p. ej., quién lo comparte), además de rasgos del propio artefacto visual. Esto refuerza la idea de que la estética no es decorativa: es parte de un régimen de autenticación distribuido entre contenido, contexto y arquitectura de la plataforma (Gehrke et al., 2025).

Una implicación teórica central es que las respuestas institucionales (p. ej., advertencias, etiquetas, alfabetización mediática) no operan en vacío: pueden producir efectos colaterales en los umbrales de confianza. Evidencia experimental ha mostrado que informar sobre deepfakes no necesariamente mejora la detección y puede aumentar la sospecha incluso ante contenidos reales, alimentando un clima donde la deslegitimación se vuelve más disponible como estrategia ("si puede ser falso, entonces lo descarto"). Este punto conecta directamente con la sección empírica del artículo, porque las "pruebas visuales" falsas funcionan tanto por su verosimilitud como por la incertidumbre que inducen (Köbis et al., 2021; Ternovski et al., 2022).

En el plano normativo, se ha argumentado que el etiquetado de contenido sintético constituye un primer paso necesario para mitigar daños epistémicos, aunque su eficacia dependa de diseño, adopción y confianza en quien etiqueta. Para este estudio, ese debate es relevante no como "solución", sino como evidencia de que la autenticidad se ha convertido en una dimensión disputada de la comunicación pública, donde estética, tecnología y gobernanza se entrelazan (Fisher, 2025).

## 3. Metodología

### *3.1. Diseño y enfoque del estudio*

El estudio se desarrolló bajo un enfoque cualitativo y un diseño descriptivo–interpretativo, no experimental y transversal, orientado a caracterizar la estética de la desinformación en noticias a partir del análisis de pruebas visuales falsas. Se adoptó un análisis multimodal de cultura visual para examinar cómo distintos recursos semióticos (imagen, tipografía, color, composición, marcas de interfaz y texto

de anclaje) se ensamblaron para producir verosimilitud y efectos de evidencia en piezas que circularon como información noticiosa.

### *3.2. Contexto y ámbito*

El análisis se situó en el ecosistema noticioso digital en español durante el periodo 30 de enero de 2024-12 de noviembre de 2025, con atención a contenidos que circularon en entornos de plataforma (X, Facebook, Instagram, TikTok, YouTube y mensajería —WhatsApp y Telegram—) y que se presentaron explícitamente como "noticia", "prueba", "captura" o "evidencia" relacionada con acontecimientos públicos. Dado que la investigación se centró en convenciones estéticas y multimodales, el ámbito se definió por géneros y formatos (capturas, montajes, contenidos sintéticos) más que por una única plataforma o medio.

### *3.3. Corpus, unidad de análisis y estrategia de muestreo*

La unidad de análisis fue cada pieza de desinformación visual entendida como un artefacto multimodal completo, compuesto por: (a) la imagen fija o video; (b) el texto incrustado en la pieza (rótulos, titulares, "sellos", marcas); y (c) el texto de acompañamiento o contexto mínimo de circulación (copy, caption, hilo, mensaje reenviado) cuando estuvo disponible y fue necesario para interpretar el anclaje.

El corpus se construyó mediante un muestreo intencional y teórico, buscando máxima variación de formatos y mecanismos de evidencialidad. Para sostener comparabilidad analítica, se equilibraron tres familias operativas de "pruebas visuales falsas": (1) capturas recontextualizadas, (2) montajes compositivos y (3) contenidos sintéticos. El corpus final estuvo compuesto por N = 60 piezas, distribuidas de forma equilibrada en tres familias (T1 = 20, T2 = 20, T3 = 20), y el tamaño se definió por saturación analítica; adicionalmente, se aseguró la comparabilidad entre familias mediante una distribución equilibrada, es decir, cuando nuevas piezas dejaron de aportar rasgos estéticos o combinaciones multimodales sustantivamente distintas.

Se incluyeron piezas que: (a) circularon como noticia o evidencia (mención explícita a "medios", "última hora", "prueba", "video", "captura", "comunicado", "informe", etc.) o imitaron convenciones noticiosas (titulares, cintillos, lower thirds, "placas" informativas); (b) fueron verificables como falsas o engañosas mediante contraste con verificaciones públicas, desmentidos institucionales o evidencia documental trazable; y (c) presentaron calidad mínima para análisis multimodal (legibilidad del texto incrustado y visibilidad de elementos de interfaz o composición). Se excluyeron materiales puramente satíricos sin circulación informativa, contenidos de opinión sin pretensión de evidencia visual y duplicados o variantes mínimas sin cambios relevantes en recursos multimodales.

### *3.4. Fuentes de datos y procedimiento de recolección y archivo*

Las piezas se recolectaron a partir de: (1) verificaciones y repositorios públicos de organizaciones de fact-checking/observatorios —Maldita.es, Newtral, Chequeado, AFP Factual, Colombiacheck y Verificat—, seleccionadas para asegurar la trazabilidad del veredicto y recuperar versiones archivadas de las piezas, (2) archivos digitales de medios y hemerotecas online para contrastar atribuciones y formatos, y (3) plataformas de circulación para recuperar contexto de publicación cuando fue imprescindible para interpretar el anclaje.

La recolección se realizó mediante: (a) identificación de piezas candidatas por palabras clave y rastreo desde verificaciones; (b) descarga/captura en su versión más completa disponible, preservando resolución suficiente para análisis de tipografía e interfaz; (c) registro de metadatos por pieza (fecha de publicación/observación, plataforma/fuente, idioma, tema, reclamo asociado, tipo de "prueba", procedencia atribuida y enlaces de referencia); y (d) archivado con identificador único (ID), almacenando el artefacto, capturas del contexto de circulación cuando correspondió y enlaces persistentes cuando estuvieron disponibles.

### *3.5 Instrumentos analíticos y categorías de codificación*

Se elaboró una matriz de codificación multimodal (codebook) para operacionalizar la "estética de la desinformación" como estrategia de evidencialidad visual. La matriz integró categorías jerárquicas con definiciones operativas y ejemplos, organizadas en seis dimensiones: (1) composición y jerarquía visual, (2) tipografía y textualidad incrustada, (3) color e indicadores de género noticioso, (4) marcas de

interfaz y procedencia, (5) anclaje texto–imagen y estructura probatoria, y (6) índices de autenticidad simulada. Las categorías se diseñaron para identificar procedimientos estéticos recurrentes y sus combinaciones, más que para cuantificar frecuencias como objetivo central.

### *3.6 Procedimiento de análisis*

El análisis se condujo en seis fases: (1) familiarización con el corpus y elaboración de memos analíticos; (2) codificación piloto de un subconjunto del corpus para depurar definiciones operativas y ajustar el codebook; (3) codificación sistemática del corpus completo, registrando segmentos multimodales y anotaciones vinculadas a cada código; (4) refinamiento iterativo de categorías y construcción de perfiles por pieza; (5) elaboración de la tipología basada en la operación dominante (recontextualización, montaje o síntesis) y el repertorio estético de evidencialidad; y (6) validación interna mediante contraste de casos típicos y casos límite para asegurar robustez conceptual.

### *3.7 Software y herramientas*

La gestión del corpus, metadatos y codificación se realizó mediante ATLAS.ti 23 y herramientas de registro estructurado (hoja de cálculo y repositorio documental). Los códigos y memos analíticos se preservaron mediante exportaciones periódicas para auditoría y respaldo del rastro de decisiones.

### *3.8 Rigor cualitativo y trazabilidad*

Para asegurar credibilidad se aplicaron: calibración del codebook mediante codificación piloto, contraste de casos típicos y atípicos, y triangulación mínima (pieza + contexto de circulación cuando estuvo disponible + fuente de verificación/contraste documental). La triangulación mínima se realizó contrastando (i) la pieza, (ii) su contexto de circulación cuando estuvo disponible y (iii) la verificación publicada o el documento/fuente original citado por el verificador (p. ej., comunicado, nota original, registro audiovisual completo). Para fortalecer la dependabilidad se mantuvo un registro de decisiones (audit trail) que documentó cambios en categorías y reglas de clasificación. Para sostener la confirmabilidad se emplearon memos de reflexividad y se preservaron ejemplos de codificación que permitieron rastrear inferencias del dato a la categoría. La consistencia del proceso se reforzó mediante revisión interna iterativa del codebook y discusión de criterios de codificación cuando surgieron ambigüedades.

### *3.9 Consideraciones éticas*

Se trabajó con contenidos disponibles públicamente o recuperados desde repositorios de verificación, aplicando minimización de daño. Cuando las piezas incluyeron datos personales o identificadores de usuarios, se anonimizó el material de trabajo y se evitó reproducir identificadores en el manuscrito, salvo en casos de interés público y difusión previa por fuentes periodísticas verificadas. El tratamiento de material sensible se gestionó mediante almacenamiento restringido y criterios editoriales de reproducción, procurando el cumplimiento de términos de servicio y normativas aplicables.

## 4. Resultados

El análisis multimodal permitió organizar las pruebas visuales falsas que circularon como contenido noticioso en una tipología de tres tipos (capturas recontextualizadas, montajes compositivos y contenidos sintéticos/manipulados). Asimismo, se sistematizaron indicios de evidencialidad (marcadores visuales que construyeron apariencia de prueba) y patrones de anclaje entre texto e imagen que orientaron la lectura hacia inferencias específicas. A continuación, se presentaron estos hallazgos en el orden: tipología → indicios de evidencialidad → anclaje texto–imagen → casos límite.

### *4.1. Tipología de pruebas visuales falsas*

La clasificación se basó en la operación principal mediante la cual cada pieza produjo una apariencia de evidencia (recontextualización, recomposición/edición o síntesis) y en su configuración multimodal (marcas de interfaz, recursos tipográficos, composición y anclaje textual). Se distinguieron los siguientes tipos analíticos (Tabla 1):

### *4.1.1. Tipo 1 (T1). Capturas recontextualizadas*

Las piezas clasificadas como T1 correspondieron a capturas de pantalla o recortes que preservaron rasgos de "registro" (p. ej., interfaz, barras de navegación, fechas aparentes, logos, estilos de publicación), pero que fueron asociadas a un reclamo que no se derivó del contexto original o que dependió de una atribución no verificable en la fuente mostrada.

#### *4.1.1.1Criterios de pertenencia:*

(a) presencia de señales de captura/interfaz (UI, timestamps, iconografía de plataforma);
(b) reclamo dependiente de atribución a una fuente o contexto específico; (c) desajuste entre el contenido visible y la interpretación propuesta (por ejemplo, recorte que elimina elementos clave o pie de foto ausente).

#### *4.1.1.2. Variación interna (subtipos operativos):*

- T1a. Capturas de redes: posts o comentarios aislados presentados como "prueba" de hechos.
- T1b. Capturas atribuidas a medios: titulares o "pantallazos" de portales/TV con formato noticioso.
- T1c. Capturas de "documentos": imágenes de supuestos comunicados, oficios o informes con membretes.

#### *4.1.1.2. Casos del corpus (ID, véase Apéndice A, Tabla A1):*

- ID-T1b-01: captura recortada de un supuesto titular con logo de medio + fecha parcial + rótulo "Última hora".
- ID-T1a-01: captura de pantalla de un post con nombre de usuario recortado + frase "confirmado" en texto añadido.

### *4.1.2. Tipo 2 (T2). Montajes compositivos*

El tipo T2 agrupó piezas que combinaron fragmentos heterogéneos mediante composición (capas, recortes, sobreimpresiones), generando una "prueba" por recomposición visual. El montaje produjo una unidad gráfica que imitó la estética de "placa noticiosa", "dossier", "collage probatorio" o "comparativo" y concentró múltiples señales de autoridad (logos, sellos, rótulos, citas).

#### *4.1.2.1. Criterios de pertenencia:*

(a) indicios de edición/ensamblaje (capas, tipografías dispares, sombras inconsistentes, bordes de recorte); (b) producción de una unidad visual que pretendió "documentar" un hecho/atribución; (c) dependencia de rótulos, sellos o títulos para estabilizar la lectura.

#### *4.1.2.2. Variación interna (subtipos operativos):*

- T2a. Placas noticiosas apócrifas: composición de "breaking news" con cintillos, lower thirds y logo.
- T2b. Fotomontaje de escena: inserción o sustitución de elementos (rostros, objetos, rótulos) para sostener atribuciones.
- T2c. Collage de múltiples "pruebas": una lámina que agrupa capturas, fotos y textos numerados como "evidencias".

#### *4.1.2.3. Casos del corpus (ID, véase Apéndice A, Tabla A1):*

- ID-T2c-01: collage con tres recortes numerados, flechas y un encabezado "PRUEBAS".
- ID-T2a-01: placa con cintillo rojo "URGENTE", tipografía bold, y logo de canal imitativo.

### *4.1.3. Tipo 3 (T3). Contenidos sintéticos o audiovisuales manipulados*

Las piezas T3 incluyeron contenidos en los que la apariencia de registro se produjo mediante la síntesis o la manipulación audiovisual (rostro/voz/movimiento, reedición de secuencia, aceleración/recorte, audio recontextualizado). En este tipo, la credibilidad se apoyó tanto en el realismo audiovisual como en recursos paratextuales que encuadraron el material como "filtrado", "exclusivo" o "prueba".

#### *4.1.3.1. Criterios de pertenencia:*

(a) soporte audiovisual o fotorrealismo de alta fidelidad; (b) señales de manipulación/síntesis reportadas por verificación externa o inconsistencias internas (p. ej., sincronía labial, artefactos); (c) reclamo dependiente de testimonio audiovisual ("se le escucha/ve decir…").

#### *4.1.3.2. Variación interna (subtipos operativos):*

- T3a. Deepfake: sustitución sintética de rostro/voz en video.
- T3b. Cheapfake: manipulación de baja complejidad (recortes, velocidad, subtítulos engañosos).
- T3c. Audio/manipulación de sincronía: audio descontextualizado o montaje de voz con imagen.

#### *4.1.3.3. Casos del corpus (ID, véase Apéndice A, Tabla A1):*

- ID-T3b-01: clip recortado con subtítulos añadidos y rótulo "CONFIRMADO".
- ID-T3a-01: video con sustitución facial + sobreimpresión de logo de medio.

**Tabla 1.** Tipología de pruebas visuales falsas y rasgos distintivos

| Tipo | Operación principal | Señales visuales predominantes | Señales de procedencia | Anclaje texto–imagen | Ejemplos del corpus (ID) |
|---|---|---|---|---|---|
| T1 Capturas recontextualizadas | Desanclaje/ recontextualización | UI visible, timestamps parciales, recortes, baja resolución por reenvío | Logos parciales, nombres de usuario recortados, "pantallazo" | Atribución ("según X"), temporalidad ("hoy"), autenticidad por "captura" | ID-T1a-01, ID-T1b-01, ID-T1c-01 |
| T2 Montajes compositivos | Recomposición/ edición | Capas, tipografías mixtas, flechas/círculos, collage, sellos | Logos apócrifos, placas noticiosas imitativas, supuestos membretes | Cierre interpretativo ("prueba de"), causalidad ("por eso"), enumeración probatoria | ID-T2a-01, ID-T2b-01, ID-T2c-01 |
| T3 Sintéticos/ manipulados | Síntesis/ manipulación audiovisual | Fotorrealismo, artefactos, subtítulos añadidos, recortes de video | "Filtrado/exclusivo", rótulos, sobreimpresión de medios | Testimonio ("se le escucha/ve"), autenticidad por registro audiovisual | ID-T3a-01, ID-T3b-01, ID-T3c-01 |

**Nota.** Los IDs remiten al registro de trazabilidad del corpus (Apéndice A, Tabla A1).

### *4.2. Indicios recurrentes de evidencialidad y "estética de prueba"*

Se sistematizaron indicios multimodales que funcionaron como indicios de evidencialidad. Estos conjuntos se derivaron de la matriz de codificación multimodal (Sección 3.5) y operaron como una síntesis transversal de códigos observados en seis dimensiones analíticas (composición y jerarquía visual; tipografía y textualidad incrustada; color e indicadores de género noticioso; marcas de interfaz y procedencia; anclaje texto–imagen y estructura probatoria; e índices de autenticidad simulada). En términos operativos, los cinco conjuntos agruparon combinaciones recurrentes de esos códigos para describir cómo se construyó la apariencia de "prueba" a través de los tres tipos (T1–T3), tal como resume la Tabla 2.

**Tabla 2.** Indicios de evidencialidad (conjuntos, recursos típicos y Casos del corpus [ID])

| Conjunto | Descripción operativa | Recursos típicos | Tipos donde se integra con facilidad | Casos del corpus (ID) |
|---|---|---|---|---|
| Procedencia/autoridad | Atribuye legitimidad mediante marcas institucionales o mediáticas | Logos, membretes, sellos "OFICIAL", firmas escaneadas | T1, T2, T3 | ID-T1c-01, ID-T2a-01 |
| Urgencia noticiosa | Construye actualidad y relevancia inmediata | Cintillos "URGENTE", "ÚLTIMA HORA", rojo/amarillo, bold | T1, T2, T3 | ID-T1b-01, ID-T2a-01 |
| Interfaz/captura | Produce apariencia de registro directo | UI, timestamps, iconos, barras, "reenviado" | T1 (central), T2 (incorporado), T3 (acompañante) | ID-T1a-01, ID-T2c-01 |
| Señalamiento probatorio | Guía la atención hacia "evidencias" | Flechas, círculos, zoom, recuadros, numeración | T2 (central), T1/T3 (apoyo) | ID-T2c-01, ID-T2b-01 |
| Documentalidad | Simula reporte técnico/administrativo | Formato expediente, tablas, encabezados, "anexos" | T1c, T2c | ID-T1c-01, ID-T2c-01 |

**Nota.** Los IDs remiten al registro de trazabilidad del corpus (Apéndice A, Tabla A1).

### *4.3. Patrones de anclaje texto–imagen en la construcción de "prueba"*

Se identificaron patrones recurrentes en la relación texto–imagen (texto incrustado y/o texto de acompañamiento) que fijaron la interpretación:

- P1. Anclaje atributivo (fuente/autoridad): el texto fijó la lectura por atribución ("según [medio/institución]"), mientras la imagen actuó como soporte de procedencia (logo, captura, membrete).
  Casos representativos (IDs del corpus): ID-T1b-01, ID-T1c-01.

- P2. Anclaje causal (detalle → conclusión): el texto extrajo una inferencia fuerte desde un detalle visual señalado (flecha/zoom), presentándolo como evidencia conclusiva del reclamo.
  Casos representativos (IDs del corpus): ID-T2b-01, ID-T2c-01.
- P3. Anclaje temporal (actualidad/ "ocurre ahora"): la pieza se apoyó en fechas aparentes, timestamps o rótulos de actualidad para instalar urgencia y contemporaneidad
  Casos representativos (IDs del corpus): ID-T1b-01, ID-T3b-01.
- P4. Anclaje testimonial (audio/imagen como declaración): el texto presentó el material como testimonio directo ("se escucha/ve decir…"), reforzado por estética audiovisual y rótulos de "filtración".
  Casos representativos (IDs del corpus): ID-T3a-01, ID-T3c-01.
  - P5. Anclaje pericial (documento → veredicto): la pieza organizó elementos en formato de "informe" y el texto concluyó mediante lenguaje técnico ("prueba", "evidencia", "peritaje"), apoyándose en documentalidad.
    Casos representativos (IDs del corpus): ID-T1c-01, ID-T2c-01.

### *4.4. Casos límite y ambigüedades clasificatorias (del corpus)*

Se registraron casos en los que una pieza combinó operaciones de más de un tipo (véase Tabla A2, Apéndice A):

- Híbrido T1↔T2 (captura dentro de collage): una captura con UI visible se integró en una lámina compuesta con rótulos, numeración y sellos, lo que introdujo rasgos de montaje (ID-H-01).
- Híbrido T2↔T3 (placa noticiosa + video manipulado): un clip manipulado se encuadró dentro de una placa "breaking news" con logo y cintillos, reforzando simultáneamente documentalidad y testimonio audiovisual (ID-H-02).
- Híbrido T1↔T3 (recontextualización audiovisual): un video real recortado y reinsertado con subtítulos y timestamp aparente se presentó como evidencia de un hecho distinto (ID-H-03).

En estos casos se utilizó el criterio de operación dominante para la clasificación: se priorizó la operación que organizó la lectura probatoria (recontextualización, montaje o síntesis/manipulación), dejando la segunda como rasgo complementario en la ficha analítica.

## 5. Discusión

### *5.1 Contribución de la tipología (T1–T3) y su alcance*

Los hallazgos permiten precisar el fenómeno de las "pruebas visuales falsas" como un régimen de verosimilitud que no depende únicamente de la manipulación técnica de la imagen, sino de operaciones semióticas y contextuales que reconfiguran su estatuto de evidencia dentro del discurso noticioso (Lecheler & Egelhofer, 2022). En esa línea, la tipología T1–T3 aporta una organización analítica consistente con la necesidad—ya identificada en el campo—de distinguir grados de sofisticación y modal richness en la desinformación visual, evitando reducir el problema a "deepfakes" como categoría totalizante (Weikmann & Lecheler, 2023).

En particular, la delimitación de T1 como recontextualización (capturas y "pruebas" fuera de contexto) dialoga con evidencia empírica que ha mostrado que, en desinformación, las imágenes suelen funcionar como "evidencia" aunque no sean falsas en sí mismas, sino mal atribuidas o reencuadradas (Brennen et al., 2021). Esta lectura refuerza la pertinencia de tratar a los cheapfakes como un núcleo explicativo de alta circulación y rendimiento persuasivo, frente a enfoques que sobredimensionan lo sintético por su novedad técnica (Gërguri, 2025). Al mismo tiempo, la coexistencia de T2 (montaje/recomposición) y T3 (síntesis/deepfakes) sugiere que el campo requiere integrar, bajo un mismo lente, la continuidad entre "trucos" de baja tecnología y generación algorítmica, entendiendo la desinformación como prácticas de decepción que articulan intención, mensaje y efectos esperados (Chadwick & Stanyer, 2022).

Desde el punto de vista analítico, el rendimiento de la tipología se apoya en una aproximación multimodal que permite describir la "prueba" como ensamblaje de recursos (imagen, texto, marcas de interfaz, composición), más que como atributo aislado del archivo visual (Wilson et al., 2023). Esta decisión es coherente con la semiótica social de la imagen, en la medida en que la función evidencial se construye por elecciones de diseño (encuadre, saliencia, vectores, modalidad) y por su inserción en

géneros y formatos reconocibles (G. R. Kress & Van Leeuwen, 2020). En esa dirección, los hallazgos matizan la idea de que lo sintético es persuasivo por "realismo" intrínseco: incluso cuando aparece T3, la eficacia depende de su integración en convenciones noticiosas y en marcos de credibilidad preexistentes (Brennen et al., 2021).

### *5.2 Indicios de evidencialidad y estética de "prueba"*

La discusión de los indicios observados sugiere que la "prueba visual" opera como estética forense popular: una gramática de trazas (flechas, recuadros, zoom, sellos, tipografías de alerta), acompañada de signos de circulación (capturas de pantalla, barras de URL, contadores, marcas de app) que funcionan como *garantes* de autenticidad. Esta lógica se alinea con investigaciones que han mostrado que, en entornos digitales, las audiencias han utilizado atajos evaluativos basados en cues de formato, presentación y procedencia para decidir qué creer y qué compartir (Ross Arguedas et al., 2024). En ese marco, las marcas de interfaz no son "ruido": pueden volverse recursos retóricos que simulan trazabilidad ("esto viene de X plataforma") y convierten la circulación en prueba ("si circula así, debe ser real").

A la vez, los hallazgos permiten situar la evidencialidad no solo en la promesa de "registro" (la fotografía como huella), sino en formas emergentes de fotorealismo algorítmico que imitan géneros fotográficos sin compartir sus condiciones indexicales. Ello es relevante porque la plausibilidad ya no depende únicamente de la fidelidad a lo real, sino de la fidelidad a convenciones visuales de lo real (Hausken, 2024). En otras palabras, la estética de prueba funciona como una tecnología cultural: un conjunto de expectativas visuales socialmente aprendidas que puede ser instrumentalizado tanto por montajes (T2) como por síntesis (T3).

### *5.3 Patrones de anclaje texto–imagen y cierre interpretativo*

Los hallazgos también muestran que la "prueba visual" rara vez se ofrece abierta a interpretación: tiende a aparecer con anclajes verbales que reducen la ambigüedad ("mira la evidencia", "aquí está la prueba", "captura filtrada") y orientan inferencias en una sola dirección. Esta dinámica es consistente con lo que la literatura ha señalado sobre el poder epistémico de las imágenes: pueden aumentar la sensación de verdad y plausibilidad incluso cuando el contenido es equívoco, precisamente porque facilitan inferencias rápidas y poco deliberativas (Newman & Schwarz, 2024). Bajo esta lógica, la imagen no "prueba" por sí misma; la prueba se produce en el acoplamiento entre imagen y enunciado que clausura el sentido.

En este punto, los hallazgos complementan estudios sobre detección de manipulaciones: aun cuando las audiencias atienden a señales visuales, la evaluación depende críticamente de señales contextuales y de encuadre informativo, lo que sugiere que el problema no se resuelve únicamente con entrenamiento perceptivo ("ver el error"), sino con alfabetización de contexto ("leer la escena mediática") (Appel & Prietzel, 2022). Además, la circulación pública del riesgo deepfake puede inducir un efecto paradójico: aumentar el escepticismo general y erosionar la confianza incluso en materiales auténticos, lo que vuelve más inestable el terreno de la evidencia visual en noticias (Ternovski et al., 2022).

### *5.4 Casos híbridos y límites de clasificación*

La presencia de casos híbridos (p. ej., captura auténtica recontextualizada y luego "reforzada" con añadidos gráficos o con micro-ediciones) sugiere que la tipología debe entenderse como operación dominante y no como categorías mutuamente excluyentes. Este hallazgo es importante porque explica cómo un mismo objeto puede migrar de T1 a T2 sin perder continuidad narrativa: se "mejora" la prueba para hacerla más legible y más convincente, sin necesariamente pasar por síntesis total. En ese contexto, el etiquetado y la acusación de "deepfake" adquieren una función política adicional: no solo alertan sobre falsedad, sino que pueden convertirse en un recurso para deslegitimar material auténtico bajo sospecha, generando incertidumbre estratégica (Hameleers & Marquart, 2023).

Este punto conecta con el problema del liar's dividend: la posibilidad de que actores acusados reencuadren evidencia verídica como "fabricada" para evadir costos reputacionales. La literatura ha documentado que la mera disponibilidad de la idea "esto puede ser fake" puede ser movilizada para sembrar duda y debilitar mecanismos de rendición de cuentas (Schiff et al., 2025). En ese escenario, la

tipología no solo ordena formas de falsificación, sino que ayuda a describir regímenes de duda: condiciones en las que la prueba visual pierde poder probatorio porque el ecosistema interpreta todo bajo sospecha.

### *5.5 Implicaciones teóricas y prácticas*

En el plano teórico, los hallazgos refuerzan la idea de que la "evidencia visual" en noticias se sostiene en una autenticidad distribuida: no es una propiedad del archivo, sino un efecto de ensamblaje entre marcas de interfaz, convenciones de género noticioso, y anclajes textuales que prometen trazabilidad. En consecuencia, la desinformación visual puede ser conceptualizada como disputa por las condiciones de legibilidad de la prueba: qué cuenta como evidencia, qué cuenta como fuente, y qué cuenta como "registro" en plataformas.

En el plano práctico, los hallazgos sugieren tres frentes de intervención. Primero, alfabetización visual orientada a verificación contextual: estrategias como la búsqueda inversa y el rastreo de procedencia han mostrado potencial para reducir la credulidad frente a visuales fuera de contexto, especialmente cuando se entrenan como hábitos de comprobación y no como destrezas "técnicas" aisladas (Qian et al., 2022). Segundo, fortalecimiento epistemológico del fact-checking: una perspectiva que describa la verificación como práctica situada—con límites, riesgos y decisiones bajo presión—permite diseñar protocolos más realistas de "prueba" y "refutación" en tiempo real (Suomalainen et al., 2025). Tercero, rediseño de rutinas y alianzas sociotécnicas: la verificación funciona como resolución de problemas en entornos plataformizados, por lo que su eficacia depende de cómo se organizan flujos, herramientas, visibilidad y colaboración con plataformas (Westlund et al., 2024).

Estas implicaciones deben leerse considerando tensiones operativas: cuando la verificación adopta lógicas de breaking news, puede incorporar costos epistemológicos (por ejemplo, simplificación de matices y aumento de errores bajo velocidad), lo que es particularmente crítico en fenómenos visuales de alta ambigüedad (Steensen et al., 2024). En paralelo, el auge de estrategias de etiquetado plantea un dilema: los avisos pueden ayudar, pero su diseño importa. La evidencia experimental reciente sugiere que etiquetas centradas en el proceso o en el daño potencial pueden producir efectos diferentes sobre creencias e intenciones de compartir, lo que obliga a decisiones informadas de implementación (Wittenberg et al., 2025). Además, las etiquetas "AI-generated" pueden inducir escepticismo generalizado incluso hacia contenidos verdaderos o humanos, si no se explica con precisión qué significa la etiqueta y qué grado de automatización supone (Altay & Gilardi, 2024). Finalmente, desde una perspectiva de gobernanza, las plataformas tienden a gestionar el problema también mediante formas de "reducción" (bajar circulación, despriorizar, friccionar) que no siempre son visibles para el público, lo que reconfigura el terreno de la evidencia al alterar qué se ve y qué no se ve (Gillespie, 2022). En suma, más que una solución única, se perfila un paquete de intervenciones coordinadas: alfabetización contextual, protocolos de verificación, diseño responsable de avisos y transparencia de moderación.

### *5.6 Limitaciones y líneas futuras*

Como en todo estudio cualitativo, los hallazgos deben interpretarse dentro de los límites de transferibilidad asociados a la delimitación del corpus, el muestreo intencional/teórico y la variabilidad de formatos y plataformas. Asimismo, la interpretación multimodal exige reflexividad: aun con reglas de codificación, existe un componente inevitablemente inferencial en la lectura de anclajes y marcas de evidencia. Futuras investigaciones pueden ampliar el alcance comparando países e idiomas, incorporando diseños longitudinales para observar la evolución de la estética de prueba, y triangulando con recepción (p. ej., entrevistas o experimentos) para evaluar cómo diferentes públicos interpretan los mismos indicios. También sería pertinente integrar un componente cuantitativo descriptivo (conteos de rasgos por tipo) sin desplazar el núcleo interpretativo, para mejorar la comunicabilidad de patrones a audiencias profesionales (periodismo y verificación).

## 6. Conclusiones

Este estudio ha caracterizado la estética de la desinformación en noticias a partir del análisis multimodal de pruebas visuales falsas, mostrando que la función probatoria se ha construido mediante la articulación de recursos visuales, paratextuales y contextuales más que por la manipulación técnica aislada de una imagen o video. En respuesta a la pregunta de investigación y al objetivo general

planteado, el análisis ha permitido organizar el fenómeno en una tipología operativa (T1–T3) —capturas recontextualizadas, montajes compositivos y contenidos sintéticos/manipulados— y ha descrito los indicios de evidencialidad que sostienen la apariencia de registro (marcas de procedencia, urgencia noticiosa, señales de interfaz/captura, señalamiento probatorio y documentalidad), así como patrones recurrentes de anclaje texto–imagen que orientan la lectura hacia conclusiones cerradas.

En términos de contribución teórica, el artículo ha aportado un lente conceptual para comprender la "prueba visual" como un ensamblaje multimodal de verosimilitud y autenticidad distribuida, y ha ofrecido criterios analíticos replicables para identificar cómo se simula evidencia en formatos que emulan convenciones periodísticas. En el plano práctico, los hallazgos han reforzado la necesidad de estrategias de verificación y alfabetización visual centradas en procedencia, contexto y diseño multimodal (y no solo en detectar "edición"), además de orientar recomendaciones para rutinas profesionales (periodismo y fact-checking) y para el diseño de avisos/etiquetas que contemplen cómo el formato y el encuadre textual condicionan la lectura de autenticidad. Estas implicaciones se mantienen transferibles en la medida en que el análisis describe mecanismos formales y retóricos observables en piezas que circulan como noticias, sin pretender generalizar más allá del alcance del corpus considerado.

Como limitaciones, el estudio ha dependido de una delimitación específica del corpus y de criterios de selección condicionados por la disponibilidad pública de materiales y por su trazabilidad; asimismo, el análisis multimodal implica decisiones interpretativas que, aunque sistematizadas mediante categorías operativas, exigen reflexividad y transparencia analítica. A partir de ello, futuras investigaciones pueden ampliar la comparación entre países/idiomas y plataformas, incorporar diseños longitudinales para observar la evolución de la estética de prueba, y triangular con estudios de recepción para examinar cómo distintos públicos interpretan los mismos indicios. También resulta pertinente complementar el enfoque con un componente cuantitativo descriptivo que estime recurrencias de rasgos o distribuciones por tipo, sin desplazar el núcleo interpretativo que permite comprender cómo se fabrica la apariencia de evidencia en el ecosistema noticioso contemporáneo.

## Referencias


Aljalabneh, A. A. (2024). Visual media literacy: Educational strategies to combat image and video disinformation on social media. *Frontiers in Communication*, *9*, 1490798. https://doi.org/10.3389/fcomm.2024.1490798

Altay, S., & Gilardi, F. (2024). People are skeptical of headlines labeled as AI-generated, even if true or human-made, because they assume full AI automation. *PNAS Nexus*, *3*(10), pgae403. https://doi.org/10.1093/pnasnexus/pgae403

Appel, M., & Prietzel, F. (2022). The detection of political deepfakes. *Journal of Computer-Mediated Communication*, *27*(4), zmac008. https://doi.org/10.1093/jcmc/zmac008

Brennen, J. S., Simon, F. M., & Nielsen, R. K. (2021). Beyond (Mis)Representation: Visuals in COVID-19 Misinformation. *The International Journal of Press/Politics*, *26*(1), 277-299. https://doi.org/10.1177/1940161220964780

Broda, E., & Strömbäck, J. (2024). Misinformation, disinformation, and fake news: Lessons from an interdisciplinary, systematic literature review. *Annals of the International Communication Association*, *48*(2), 139-166. https://doi.org/10.1080/23808985.2024.2323736

Chadwick, A., & Stanyer, J. (2022). Deception as a Bridging Concept in the Study of Disinformation, Misinformation, and Misperceptions: Toward a Holistic Framework. *Communication Theory*, *32*(1), 1-24. https://doi.org/10.1093/ct/qtab019

Fisher, S. A. (2025). Something AI Should Tell You – The Case for Labelling Synthetic Content. *Journal of Applied Philosophy*, *42*(1), 272-286. https://doi.org/10.1111/japp.12758

Gehrke, M., Eggers, J., De Vreese, C., & Hopmann, D. N. (2025). What Makes News (Seem) Authentic on Social Media? Indicators from a Qualitative Study of Young Adults. *Digital Journalism*, *13*(8), 1503-1522. https://doi.org/10.1080/21670811.2024.2399622

Gërguri, D. (2025). Cheapfakes. En A. Nai, M. Grömping, & D. Wirz (Eds.), *Elgar Encyclopedia of Political Communication* (pp. 192-195). Edward Elgar Publishing. https://doi.org/10.4337/9781035301447.vol1.00053

Gillespie, T. (2022). Do Not Recommend? Reduction as a Form of Content Moderation. *Social Media + Society*, *8*(3), 20563051221117552. https://doi.org/10.1177/20563051221117552

Habgood-Coote, J. (2023). Deepfakes and the epistemic apocalypse. *Synthese*, *201*(3), 103. https://doi.org/10.1007/s11229-023-04097-3

Hameleers, M. (2024). Cheap Versus Deep Manipulation: The Effects of Cheapfakes Versus Deepfakes in a Political Setting. *International Journal of Public Opinion Research*, *36*(1), edae004. https://doi.org/10.1093/ijpor/edae004

Hameleers, M. (2025). The Nature of Visual Disinformation Online: A Qualitative Content Analysis of Alternative and Social Media in the Netherlands. *Political Communication*, *42*(1), 108-126. https://doi.org/10.1080/10584609.2024.2354389

Hameleers, M., & Marquart, F. (2023). It's Nothing but a Deepfake! The Effects of Misinformation and Deepfake Labels Delegitimizing an Authentic Political Speech. *International Journal of Communication*, *17*, 21-21.

Hameleers, M., Powell, T. E., Van Der Meer, T. G. L. A., & Bos, L. (2020). A Picture Paints a Thousand Lies? The Effects and Mechanisms of Multimodal Disinformation and Rebuttals Disseminated via Social Media. *Political Communication*, *37*(2), 281-301. https://doi.org/10.1080/10584609.2019.1674979

Hausken, L. (2024). Photorealism versus photography. AI-generated depiction in the age of visual disinformation. *Journal of Aesthetics & Culture*, *16*(1), 2340787. https://doi.org/10.1080/20004214.2024.2340787

Heley, K., Gaysynsky, A., & King, A. J. (2022). Missing the Bigger Picture: The Need for More Research on Visual Health Misinformation. *Science Communication*, *44*(4), 514-527. https://doi.org/10.1177/10755470221113833

Köbis, N. C., Doležalová, B., & Soraperra, I. (2021). Fooled twice: People cannot detect deepfakes but think they can. *iScience*, *24*(11), 103364. https://doi.org/10.1016/j.isci.2021.103364

Kress, G. R., & Van Leeuwen, T. (2020). *Reading images: The grammar of visual design* (Third edition). Routledge.

Kress, G., & Van Leeuwen, T. (2020). *Reading Images: The Grammar of Visual Design* (3.ª ed.). Routledge. https://doi.org/10.4324/9781003099857

Lecheler, S., & Egelhofer, J. L. (2022). Disinformation, Misinformation, and Fake News: Understanding the Supply Side. En *Knowledge Resistance in High-Choice Information Environments*. Routledge.

Mirsky, Y., & Lee, W. (2022). The Creation and Detection of Deepfakes: A Survey. *ACM Computing Surveys*, *54*(1), 1-41. https://doi.org/10.1145/3425780

Newman, E. J., & Schwarz, N. (2024). Misinformed by images: How images influence perceptions of truth and what can be done about it. *Current Opinion in Psychology*, *56*, 101778. https://doi.org/10.1016/j.copsyc.2023.101778

Peng, Y., Lu, Y., & Shen, C. (2023). An Agenda for Studying Credibility Perceptions of Visual Misinformation. *Political Communication*, *40*(2), 225-237. https://doi.org/10.1080/10584609.2023.2175398

Pulm, C., Gast, A., & Rummel, J. (2024). A Picture Corrects a Thousand Words – The Effect of Photos on Veracity Feedback. *Consciousness and Cognition*, *125*, 103758. https://doi.org/10.1016/j.concog.2024.103758

Qian, S., Shen, C., & Zhang, J. (2022). Fighting cheapfakes: Using a digital media literacy intervention to motivate reverse search of out-of-context visual misinformation. *Journal of Computer-Mediated Communication*, *28*(1), zmac024. https://doi.org/10.1093/jcmc/zmac024

Rauchfleisch, A., Vogler, D., & De Seta, G. (2025). Deepfakes or Synthetic Media? The Effect of Euphemisms for Labeling Technology on Risk and Benefit Perceptions. *Social Media + Society*, *11*(3), 20563051251350975. https://doi.org/10.1177/20563051251350975

Ross Arguedas, A. A., Badrinathan, S., Mont'Alverne, C., Toff, B., Fletcher, R., & Nielsen, R. K. (2024). Shortcuts to trust: Relying on cues to judge online news from unfamiliar sources on digital platforms. *Journalism*, *25*(6), 1207-1229. https://doi.org/10.1177/14648849231194485

Schiff, K. J., Schiff, D. S., & Bueno, N. S. (2025). The Liar's Dividend: Can Politicians Claim Misinformation to Evade Accountability? *American Political Science Review*, *119*(1), 71-90. https://doi.org/10.1017/S0003055423001454

Shen, M., Luo, Z., & Liao, Y. (2024). Comprehensive Out-of-context Misinformation Detection via Global Information Enhancement. *Proceedings of the 2024 2nd International Conference on Advances in Artificial Intelligence and Applications*, 138-142. https://doi.org/10.1145/3712623.3712660

Steensen, S., Kalsnes, B., & Westlund, O. (2024). The limits of live fact-checking: Epistemological consequences of introducing a breaking news logic to political fact-checking. *New Media & Society*, *26*(11), 6347-6365. https://doi.org/10.1177/14614448231151436

Suomalainen, K., Nykänen, N., Seeck, H., Kim, Y., & McPherson, E. (2025). Fact-Checking in Journalism: An Epistemological Framework. *Journalism Studies*, *26*(10), 1129-1149. https://doi.org/10.1080/1461670X.2025.2492729

Ternovski, J., Kalla, J., & Aronow, P. (2022). Negative Consequences of Informing Voters about Deepfakes: Evidence from Two Survey Experiments. *Journal of Online Trust and Safety*, *1*(2). https://doi.org/10.54501/jots.v1i2.28

Tonglet, J., Thiem, G., & Gurevych, I. (2025). COVE: COntext and VEracity prediction for out-of-context images. *Proceedings of the 2025 Conference of the Nations of the Americas Chapter of the Association for Computational Linguistics: Human Language Technologies (Volume 1: Long Papers)*, 2029-2049. https://doi.org/10.18653/v1/2025.naacl-long.102

Vaccari, C., & Chadwick, A. (2020). Deepfakes and Disinformation: Exploring the Impact of Synthetic Political Video on Deception, Uncertainty, and Trust in News. *Social Media + Society*, *6*(1), 2056305120903408. https://doi.org/10.1177/2056305120903408

Vinhas, O., & Bastos, M. (2022). Fact-Checking Misinformation: Eight Notes on Consensus Reality. *Journalism Studies*, *23*(4), 448-468. https://doi.org/10.1080/1461670X.2022.2031259

Vinhas, O., & Bastos, M. (2025). The WEIRD governance of fact-checking and the politics of content moderation. *New Media & Society*, *27*(5), 2768-2787. https://doi.org/10.1177/14614448231213942

Weikmann, T., & Lecheler, S. (2023). Visual disinformation in a digital age: A literature synthesis and research agenda. *New Media & Society*, *25*(12), 3696-3713. https://doi.org/10.1177/14614448221141648

Weikmann, T., & Lecheler, S. (2024). Cutting through the Hype: Understanding the Implications of Deepfakes for the Fact-Checking Actor-Network. *Digital Journalism*, *12*(10), 1505-1522. https://doi.org/10.1080/21670811.2023.2194665

Westlund, O., Belair-Gagnon, V., Graves, L., Larsen, R., & Steensen, S. (2024). What Is the Problem with Misinformation? Fact-checking as a Sociotechnical and Problem-Solving Practice. *Journalism Studies*, *25*(8), 898-918. https://doi.org/10.1080/1461670X.2024.2357316

Wilson, A., Wilkes, S., Teramoto, Y., & Hale, S. (2023). Multimodal analysis of disinformation and misinformation. *Royal Society Open Science*, *10*(12), 230964. https://doi.org/10.1098/rsos.230964

Wittenberg, C., Epstein, Z., Péloquin-Skulski, G., Berinsky, A. J., & Rand, D. G. (2025). Labeling AI-generated media online. *PNAS Nexus*, *4*(6), pgaf170. https://doi.org/10.1093/pnasnexus/pgaf170

Zeng, J., & Brennen, S. B. (2023). Misinformation. *Internet Policy Review*, *12*(4). https://doi.org/10.14763/2023.4.1725

# Apéndice A

**Tabla A1.** Registro de trazabilidad del corpus (N = 60): tipología (T1–T3), metadatos y rasgo estético dominante

| ID | Tipo/Subtipo | Reclamo/tema (1 línea) | Rasgo estético dominante (1 línea) | Entorno | País/ámbito | Fecha (obs.) | Fuente (FC) | Token verif/contraste | Token preservación | Material | Nota ética |
|---|---|---|---|---|---|---|---|---|---|---|---|
| ID-T1a-01 | T1a | Captura de post recontextualizado como "confirmación" de hecho público | UI de red visible + recorte de usuario + texto añadido tipo "confirmado" | X | LatAm | 14/02/2024 | FC4 | VERIF_FC4_T1a01 | ARCH_T1a01 | img | usuario anon. |
| ID-T1a-02 | T1a | Captura de comentario presentada como "prueba" de cierre/medida urgente | Captura de comentario aislado + resaltado/corte selectivo + rótulo de urgencia incrustado | FB | MX | 9/03/2024 | FC1 | VERIF_FC1_T1a02 | ARCH_T1a02 | img | usuario anon. |
| ID-T1a-03 | T1a | Captura de hilo atribuida a "fuente interna" para sostener denuncia sin respaldo | UI de hilo + fragmentación por recorte + anclaje textual "fuente interna" | X | ES | 22/04/2024 | FC2 | VERIF_FC2_T1a03 | ARCH_T1a03 | img | usuario anon. |
| ID-T1a-04 | T1a | Captura de mensaje reenviado usada como evidencia de alarma sanitaria | Indicador "reenviado" /chat + timestamp aparente + mayúsculas alarmistas | WA | PE | 3/06/2024 | FC4 | VERIF_FC4_T1a04 | ARCH_T1a04 | img | datos sensibles anon. |
| ID-T1a-05 | T1a | Captura de story recontextualizada como "última hora" sobre evento masivo | Estética de story + sticker/overlay de "última hora" + recorte sin contexto | IG | CO | 18/07/2024 | FC5 | VERIF_FC5_T1a05 | ARCH_T1a05 | img | usuario anon. |
| ID-T1a-06 | T1a | Captura de post con recorte del nombre para sostener acusación a autoridad | Recorte de nombre/autor + UI parcial + texto acusatorio superpuesto | FB | CL | 1/09/2024 | FC4 | VERIF_FC4_T1a06 | ARCH_T1a06 | img | usuario anon. |
| ID-T1a-07 | T1a | Captura de chat presentada como "filtración" de decisión institucional | Captura de chat + marcas de "filtrado" /reenviado + recorte de participantes | TG | ES | 11/01/2025 | FC1 | VERIF_FC1_T1a07 | ARCH_T1a07 | img | datos personales anon. |
| ID-T1b-01 | T1b | Pantallazo de supuesto titular con logo y fecha parcial ("Última hora") | Logo de medio + fecha parcial/timestamp + rótulo "Última hora" estilo alerta | WEB | ES | 28/02/2024 | FC7 | CONT_FC7_T1b01 | ARCH_T1b01 | img | sin datos personales |
| ID-T1b-02 | T1b | Captura de "placa" atribuida a canal TV con cintillo de urgencia | Cintillo urgente + lower third imitativo + logo/bug de canal recortado | FB | LatAm | 6/05/2024 | FC4 | VERIF_FC4_T1b02 | ARCH_T1b02 | img | sin datos personales |
| ID-T1b-03 | T1b | Pantallazo de portal noticioso atribuido a medio nacional (titular apócrifo) | Encabezado de portal + tipografía de titular + URL/elementos clave recortados | X | MX | 12/08/2024 | FC2 | VERIF_FC2_T1b03 | ARCH_T1b03 | img | sin datos personales |
| ID-T1b-04 | T1b | Captura de "alerta" atribuida a medio: desanclaje temporal ("hoy") | Rótulo temporal "hoy/ahora" + captura comprimida + atribución a medio sin trazabilidad | WA | CO | 3/10/2024 | FC5 | VERIF_FC5_T1b04 | ARCH_T1b04 | img | sin datos personales |
| ID-T1b-05 | T1b | Pantallazo de nota atribuida a medio regional con URL recortada | Captura de nota + URL/fuente recortada + layout noticioso parcial | TG | AR | 17/02/2025 | FC3 | VERIF_FC3_T1b05 | ARCH_T1b05 | img | sin datos personales |
| ID-T1b-06 | T1b | Captura de "breaking" atribuida a noticiero con lower third imitativo | Lower third + "breaking" + paleta rojo/amarillo + logo imitativo | FB | PE | 9/05/2025 | FC4 | VERIF_FC4_T1b06 | ARCH_T1b06 | img | sin datos personales |
| ID-T1b-07 | T1b | Pantallazo de supuesto ranking/dato "oficial" atribuido a medio | Gráfico/"ranking" en formato noticia + sello "oficial" textual + marca de medio parcial | WEB | CL | 21/09/2025 | FC4 | VERIF_FC4_T1b07 | ARCH_T1b07 | img | sin datos personales |
| ID-T1c-01 | T1c | Supuesto comunicado con membrete y firma escaneada para sostener medida estatal | Membrete institucional + firma escaneada + sello "OFICIAL/URGENTE" | WA | ES | 15/03/2024 | FC8 | CONT_FC8_T1c01 | ARCH_T1c01 | img | datos sensibles anon. |
| ID-T1c-02 | T1c | Captura de "oficio" con sello "OFICIAL" y numeración irregular | Sello "OFICIAL" + numeración/folio + tipografía inconsistente tipo documento | TG | CO | 27/06/2024 | FC5 | VERIF_FC5_T1c02 | ARCH_T1c02 | img | datos sensibles anon. |
| ID-T1c-03 | T1c | Supuesto informe con logotipo institucional y formato "anexo" | Logotipo institucional + encabezado formal + estructura de "anexo" | FB | MX | 8/11/2024 | FC4 | VERIF_FC4_T1c03 | ARCH_T1c03 | img | datos sensibles anon. |
| ID-T1c-04 | T1c | Captura de carta atribuida a ministerio con firma sin trazabilidad | Carta formal + firma/"sello" sin rastro + fecha/membrete ambiguos | WEB | PE | 2/03/2025 | FC4 | VERIF_FC4_T1c04 | ARCH_T1c04 | img | datos sensibles anon. |
| ID-T1c-05 | T1c | Supuesto "boletín" con encabezado institucional y fecha dudosa | Encabezado institucional + fecha dudosa/formatos mezclados + tono administrativo | X | AR | 14/06/2025 | FC3 | VERIF_FC3_T1c05 | ARCH_T1c05 | img | datos sensibles anon. |
| ID-T1c-06 | T1c | Captura de "resolución" con sellos múltiples y tipografía inconsistente | Multiplicidad de sellos + tipografía dispareja + apariencia de resolución oficial | FB | CL | 5/10/2025 | FC4 | VERIF_FC4_T1c06 | ARCH_T1c06 | img | datos sensibles anon. |
| ID-T2a-01 | T2a | Placa "URGENTE" apócrifa con logo imitativo y lower third | Placa TV imitativa + cintillo "URGENTE" + logo/bug apócrifo | FB | LatAm | 30/01/2024 | FC4 | VERIF_FC4_T2a01 | ARCH_T2a01 | img | sin datos personales |

| | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| ID-T2a-02 | T2a | Placa tipo noticiero con paleta roja/amarilla y rótulo "CONFIRMADO" | Alto contraste rojo/amarillo + tipografía bold + rótulo "CONFIRMADO" | X | ES | 7/04/2024 | FC1 | VERIF_FC1_T2a02 | ARCH_T2a02 | img | sin datos personales |
| ID-T2a-03 | T2a | Placa breaking news con tipografía bold y marca de canal imitativa | "Breaking news" + marca de canal imitativa + lower third con datos cerrados | IG | MX | 2/07/2024 | FC2 | VERIF_FC2_T2a03 | ARCH_T2a03 | img | sin datos personales |
| ID-T2a-04 | T2a | Placa "ALERTA" con fecha insertada y barra inferior estilo TV | Barra inferior estilo TV + fecha insertada + rótulo "ALERTA" | FB | CO | 19/09/2024 | FC5 | VERIF_FC5_T2a04 | ARCH_T2a04 | img | sin datos personales |
| ID-T2a-05 | T2a | Placa de "última hora" con logo apócrifo y supuesta cita de autoridad | "Última hora" + cita atribuida + logo apócrifo + composición tipo noticiero | WA | PE | 10/12/2024 | FC4 | VERIF_FC4_T2a05 | ARCH_T2a05 | img | sin datos personales |
| ID-T2a-06 | T2a | Placa noticiosa con rótulo "OFICIAL" y supuesta fuente institucional | Sello "OFICIAL" + fuente institucional invocada + estética de noticiero | TG | AR | 18/04/2025 | FC3 | VERIF_FC3_T2a06 | ARCH_T2a06 | img | sin datos personales |
| ID-T2a-07 | T2a | Placa imitativa con recuadro de "datos" y sello de verificación falso | Recuadro "datos" + sello/ícono de verificación falso + tipografías mixtas | X | CL | 27/08/2025 | FC4 | VERIF_FC4_T2a07 | ARCH_T2a07 | img | sin datos personales |
| ID-T2b-01 | T2b | Fotomontaje de escena: inserción de objeto/rotulado para sostener acusación | Inserción de elemento + sombras/bordes inconsistentes + rótulo explicativo | FB | ES | 5/02/2024 | FC1 | VERIF_FC1_T2b01 | ARCH_T2b01 | img | rostros anon. |
| ID-T2b-02 | T2b | Fotomontaje: sustitución de señalética con texto incrustado "prueba" | Sustitución de texto/señalética + tipografía no integrada + marca "prueba" | IG | MX | 21/05/2024 | FC4 | VERIF_FC4_T2b02 | ARCH_T2b02 | img | rostros anon. |
| ID-T2b-03 | T2b | Montaje de foto con rótulo y flecha que "demuestra" supuesto detalle | Flecha/círculo señalador + rótulo conclusivo + recorte focalizado | FB | CO | 25/08/2024 | FC5 | VERIF_FC5_T2b03 | ARCH_T2b03 | img | rostros anon. |
| ID-T2b-04 | T2b | Montaje compositivo con tipografías dispares y sombras inconsistentes | Capas visibles + tipografías dispares + sombras inconsistentes | X | PE | 28/10/2024 | FC4 | VERIF_FC4_T2b04 | ARCH_T2b04 | img | rostros anon. |
| ID-T2b-05 | T2b | Montaje de evento público con inserción de logo institucional apócrifo | Inserción de logo apócrifo + etiqueta institucional + composición de "prueba" | WEB | AR | 2/02/2025 | FC3 | VERIF_FC3_T2b05 | ARCH_T2b05 | img | rostros anon. |
| ID-T2b-06 | T2b | Montaje de imagen con recorte visible y subtítulo engañoso incrustado | Bordes de recorte + subtítulo incrustado engañoso + anclaje textual fuerte | TG | CL | 6/06/2025 | FC4 | VERIF_FC4_T2b06 | ARCH_T2b06 | img | rostros anon. |
| ID-T2b-07 | T2b | Montaje de "comparativa" antes/después con rótulos y círculos probatorios | Comparativa antes/después + círculos probatorios + rótulos de conclusión | FB | LatAm | 12/11/2025 | FC4 | VERIF_FC4_T2b07 | ARCH_T2b07 | img | rostros anon. |
| ID-T2c-01 | T2c | Collage probatorio: tres recortes numerados + flechas + encabezado "PRUEBAS" | Collage con numeración + flechas + encabezado "PRUEBAS" (dossier) | WA | ES | 29/03/2024 | FC2 | VERIF_FC2_T2c01 | ARCH_T2c01 | img | usuario anon. |
| ID-T2c-02 | T2c | Collage: capturas + foto + "sellos" para estabilizar interpretación causal | Collage mixto + sellos/etiquetas + texto causal que fija lectura | FB | MX | 11/06/2024 | FC4 | VERIF_FC4_T2c02 | ARCH_T2c02 | img | usuario anon. |
| ID-T2c-03 | T2c | Collage tipo "dossier" con logos múltiples y citas atribuidas sin fuente | Logos múltiples + citas sin fuente + diseño tipo dossier probatorio | X | CO | 7/09/2024 | FC5 | VERIF_FC5_T2c03 | ARCH_T2c03 | img | usuario anon. |
| ID-T2c-04 | T2c | Collage con recuadros y numeración que organiza "evidencias" de un reclamo | Recuadros + numeración + jerarquía visual tipo "expediente" | WEB | PE | 24/01/2025 | FC4 | VERIF_FC4_T2c04 | ARCH_T2c04 | img | usuario anon. |
| ID-T2c-05 | T2c | Collage de "documentos" con membretes y resaltados tipo pericial | Membretes + resaltado/zoom + estructura pericial (anexos) | TG | AR | 28/05/2025 | FC3 | VERIF_FC3_T2c05 | ARCH_T2c05 | img | datos sensibles anon. |
| ID-T2c-06 | T2c | Collage con capturas de interfaz + rótulo "OFICIAL" + conclusión cerrada | Capturas UI integradas + sello "OFICIAL" + conclusión textual final | FB | CL | 2/09/2025 | FC4 | VERIF_FC4_T2c06 | ARCH_T2c06 | img | usuario anon. |
| ID-T3a-01 | T3a | Deepfake: rostro/voz de figura pública en declaración "exclusiva" | Sustitución facial/voz + artefactos de sincronía + encuadre "exclusivo" | TT | LatAm | 19/02/2024 | FC4 | VERIF_FC4_T3a01 | ARCH_T3a01 | video | rostro sensible (uso restringido) |
| ID-T3a-02 | T3a | Deepfake: video con sustitución facial + logo de medio sobreimpreso | Sustitución facial + logo de medio sobreimpreso + lower third falso | FB | ES | 14/05/2024 | FC1 | VERIF_FC1_T3a02 | ARCH_T3a02 | video | rostro sensible (uso restringido) |
| ID-T3a-03 | T3a | Deepfake: declaración atribuida a autoridad con artefactos de sincronía | Desajuste labial + textura/artefactos + anclaje "dijo/confirmó" | X | MX | 4/08/2024 | FC2 | VERIF_FC2_T3a03 | ARCH_T3a03 | video | rostro sensible (uso restringido) |
| ID-T3a-04 | T3a | Deepfake: "anuncio" político con rótulo "CONFIRMADO" y baja trazabilidad | Rótulo "CONFIRMADO" + estética noticiosa + señales de síntesis | FB | CO | 16/10/2024 | FC5 | VERIF_FC5_T3a04 | ARCH_T3a04 | video | rostro sensible (uso restringido) |

| | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| ID-T3a-05 | T3a | Deepfake: pieza presentada como "filtración" con marcas de edición | Marco "filtrado" + cortes/edición visibles + inconsistencias faciales | TT | PE | 26/02/2025 | FC4 | VERIF_FC4_T3a05 | ARCH_T3a05 | video | rostro sensible (uso restringido) |
| ID-T3a-06 | T3a | Deepfake: sustitución de voz con plano fijo y sello "OFICIAL" | Voz sintética + plano fijo + sello "OFICIAL" para autoridad simulada | TG | AR | 30/06/2025 | FC3 | VERIF_FC3_T3a06 | ARCH_T3a06 | video | rostro sensible (uso restringido) |
| ID-T3a-07 | T3a | Deepfake: clip con sobreimpresión de canal imitativo y audio sintético | Bug de canal imitativo + audio sintético + artefactos de rostro | X | CL | 19/10/2025 | FC4 | VERIF_FC4_T3a07 | ARCH_T3a07 | video | rostro sensible (uso restringido) |
| ID-T3b-01 | T3b | Cheapfake: clip recortado con subtítulos añadidos que cambian el sentido | Recorte selectivo + subtítulos añadidos + rótulo de certeza/confirmación | YT | ES | 4/03/2024 | FC1 | VERIF_FC1_T3b01 | ARCH_T3b01 | video | rostros anon. |
| ID-T3b-02 | T3b | Cheapfake: aceleración/recorte + rótulo "última hora" para generar alarma | Velocidad alterada + "última hora" + alto contraste tipo alerta | FB | MX | 1/06/2024 | FC4 | VERIF_FC4_T3b02 | ARCH_T3b02 | video | rostros anon. |
| ID-T3b-03 | T3b | Cheapfake: inserción de subtítulos engañosos sobre video real | Subtítulos engañosos + recorte contextual + anclaje causal en texto | TT | CO | 28/09/2024 | FC5 | VERIF_FC5_T3b03 | ARCH_T3b03 | video | rostros anon. |
| ID-T3b-04 | T3b | Cheapfake: clip fuera de contexto con timestamp aparente y conclusión textual | Timestamp aparente + clip fuera de contexto + conclusión textual cerrada | X | PE | 22/12/2024 | FC4 | VERIF_FC4_T3b04 | ARCH_T3b04 | video | rostros anon. |
| ID-T3b-05 | T3b | Cheapfake: reedición de secuencia + rótulo "PRUEBA" en pantalla | Reedición + rótulo "PRUEBA" + inserción de texto probatorio | FB | AR | 20/03/2025 | FC3 | VERIF_FC3_T3b05 | ARCH_T3b05 | video | rostros anon. |
| ID-T3b-06 | T3b | Cheapfake: montaje de subtítulos + logo de medio para estabilizar lectura | Logo de medio + subtítulos + estética de noticia para fijar interpretación | IG | CL | 7/07/2025 | FC4 | VERIF_FC4_T3b06 | ARCH_T3b06 | video | rostros anon. |
| ID-T3b-07 | T3b | Cheapfake: recorte selectivo + texto de anclaje "se confirma" | Texto de anclaje "se confirma" + recorte selectivo + urgencia visual | TT | LatAm | 3/11/2025 | FC4 | VERIF_FC4_T3b07 | ARCH_T3b07 | video | rostros anon. |
| ID-T3c-01 | T3c | Audio descontextualizado sobre imagen fija presentado como "testimonio" | Audio recontextualizado + imagen fija + rótulo "se escucha/filtrado" | WA | ES | 18/04/2024 | FC2 | VERIF_FC2_T3c01 | ARCH_T3c01 | audio+img | voz anon. |
| ID-T3c-02 | T3c | Audio atribuido a autoridad montado sobre video de archivo | Audio atribuido + video de archivo + texto de atribución a autoridad | FB | MX | 23/07/2024 | FC4 | VERIF_FC4_T3c02 | ARCH_T3c02 | video | voz anon. |
| ID-T3c-03 | T3c | Desfase audio/imagen + rótulo "filtrado" como marco de autenticidad | Desfase audio/imagen + rótulo "filtrado" + estética de evidencia | X | CO | 7/10/2024 | FC5 | VERIF_FC5_T3c03 | ARCH_T3c03 | video | voz anon. |
| ID-T3c-04 | T3c | Audio recortado + capturas de interfaz para reforzar "prueba" | Capturas de interfaz + audio recortado + timestamp/controles visibles | TG | PE | 8/01/2025 | FC4 | VERIF_FC4_T3c04 | ARCH_T3c04 | video | voz anon. |
| ID-T3c-05 | T3c | Audio recontextualizado con subtítulo incrustado y sello "OFICIAL" | Subtítulo incrustado + sello "OFICIAL" + atribución cerrada | FB | AR | 4/05/2025 | FC3 | VERIF_FC3_T3c05 | ARCH_T3c05 | video | voz anon. |
| ID-T3c-06 | T3c | Montaje de voz sobre imagen de noticiero para simular declaración | Voz montada + fondo noticiero + logo/placa para legitimidad | WEB | CL | 29/09/2025 | FC4 | VERIF_FC4_T3c06 | ARCH_T3c06 | video | voz anon. |

**Tabla A2.** Casos límite/híbridos (ID-H-01 a ID-H-03): operación dominante ↔ secundaria y metadatos de trazabilidad

| ID | Híbrido (operación dominante ↔ secundaria) | Reclamo/tema (1 línea) | Entorno | País/ámbito | Fecha (obs.) | Fuente (FC) | Token verif/contraste | Token preservación | Material | Nota ética |
|---|---|---|---|---|---|---|---|---|---|---|
| ID-H-01 | T2 ↔ T1 (montaje dominante; captura incorporada) | Captura con UI integrada en lámina compuesta con rótulos/sellos para sostener inferencia cerrada | FB | LatAm | 15/08/2024 | FC4 | VERIF_FC4_H01 | ARCH_H01 | img | usuario anon. / posible dato sensible anon. |
| ID-H-02 | T2 ↔ T3 (placa dominante; audiovisual manipulado) | Clip manipulado encuadrado como "breaking news" con logo/cintillos para reforzar testimonio y urgencia | X | ES | 11/03/2025 | FC1 | VERIF_FC1_H02 | ARCH_H02 | video | rostros anon. / uso restringido si figura pública |
| ID-H-03 | T1 ↔ T3 (recontextualización dominante; edición audiovisual) | Video real recortado y reinsertado con subtítulos/timestamp aparente para atribuir hecho distinto | WA | PE | 12/11/2025 | FC4 | VERIF_FC4_H03 | ARCH_H03 | video | minimizar daño; no reproducir identificadores |