VISUAL REVIEW | Vol. 18, No. 2, 2026 | ISSN 2695-9631
International Visual Culture Review / Revista Internacional de Cultura Visual
https://doi.org/10.62161/revvisual.v18.5930

ÉTICA, IMAGEN PERSONAL Y DESINFORMACIÓN EN LA ERA DE LOS *DEEPFAKES*

ÁNGEL FERNÁNDEZ FERNÁNDEZ [1]
amfernandez@thecoreschool.com
IREIDE MARTÍNEZ DE BARTOLOME RINCÓN [2]
evaireide.martinez@universidadunie.com
BORJA MORGADO AGUIRRE [3]
morgado@um.es

[1] The Core School / UNIE Universidad
[2] UNIE Universidad
[3] Universidad de Murcia

PALABRAS CLAVE

Deepfake
Inteligencia Artificial
Ética
Desinformación
Imagen personal

RESUMEN

La creciente relevancia de la cultura visual digital plantea grandes interrogantes relacionados con la curación de imágenes, el respeto a los derechos de imagen y la protección de la privacidad. En este contexto, el desarrollo de la tecnología deepfakes agrava estos problemas al permitir una manipulación audiovisual sin precedentes. Este artículo analiza las implicaciones éticas, legales y sociales de la creación y difusión de deepfakes, enfatizando la problemática de la falsificación de identidades. Asimismo, examina las medidas de control y regulación implementadas por las plataformas digitales para detectar y limitar la distribución de este tipo de contenido, apuntando hacia la necesidad de pautas éticas claras. También se aborda la capacidad de los deepfakes para reforzar sesgos y narrativas discriminatorias, deteriorando la confianza en la información visual perpetuando estereotipos dañinos en el imaginario colectivo.

Recibido: 27/ 06 / 2025
Aceptado: 07/ 09 / 2025

1. Introducción

La presente investigación tiene como objetivo explorar las repercusiones éticas y legales derivadas de la creación y de la difusión de *deepfakes*, analizando la responsabilidad de quienes comparten, editan o recopilan este tipo de imágenes y proponiendo herramientas y lineamientos éticos que protejan la privacidad y la equidad en un entorno cada vez más afectado por la manipulación visual. El análisis se centrará en dos áreas principales: el impacto de los *deepfakes* en los derechos individuales y la imagen y la capacidad de esta tecnología para reforzar sesgos y narrativas discriminatorias. Los *deepfakes* representan una amenaza sin precedentes para la confianza pública, ya que cuestionan la validez de las imágenes y fomentan la desinformación a gran escala. Es urgente abordar el problema desde una perspectiva interdisciplinaria que combine ética, derecho y tecnología (Chesney y Citron, 2019).

1.1. Antecedentes

La preeminencia cada vez mayor de la cultura visual en los entornos digitales ha transformado profundamente las dinámicas de comunicación y representación contemporáneas. Actualmente, las imágenes, especialmente la fotografía y el vídeo, han alcanzado un protagonismo comunicacional sin precedentes, mediatizando nuestra mirada y condicionando el modo en que interpretamos la realidad. Desde finales del siglo XX, cuando los medios digitales todavía se encontraban en un momento incipiente, Baudrillard (1999) ya hacía referencia a la *promiscuidad* y *ubicuidad* de las imágenes, destacando su capacidad de contaminación cultural y su viralidad. Hoy vivimos en un estado de hipervisualidad (Buxó, 1999), una suerte de cultura visual digital en la que las imágenes ejercen una influencia simbólica excepcional. En este nuevo imperio de lo visual, la imagen se ha consolidado como el principal soporte comunicativo, permeando y transformando todos los aspectos de nuestra experiencia (Fernández, 2017; Westerlund, 2019).

Este nuevo modelo de interpretación de la imagen plantea profundos desafíos éticos que impactan en nuestra percepción de la realidad y nos obligan a reconsiderar aspectos tan diversos como el papel de las imágenes en la construcción de nuestra identidad, el control de la privacidad en los entornos digitales o la gestión de la desinformación (Chesney y Citron, 2019; Vaccari y Chadwick, 2020). Tales retos se han intensificado tras los últimos avances de la Inteligencia Artificial Generativa (IAG), en particular, la irrupción de los *deepfakes,* imágenes falsas aunque extremadamente realistas, generadas mediante sistemas de IA (Ajder et al., 2019; Bañuelos, 2022), ha permitido copiar, iterar o sustituir la imagen y la voz de una persona por las de otra, a través de redes neuronales de aprendizaje profundo (Bode et al., 2021).

Actualmente, la rápida evolución de la tecnología que sustenta los *deepfakes* está desafiando las nociones tradicionales de autenticidad y verosimilitud, minando la confianza en el carácter referencial de las imágenes. Aunque las capacidades de la inteligencia artificial (IA) aún no están completamente definidas y la regulación sobre sus límites es todavía escasa, es evidente que los *deepfakes* están teniendo un impacto significativo en la estructura misma de los medios digitales. De hecho, a pesar de que pueden ser empleados con fines lícitos y ofrecen aplicaciones innovadoras en ámbitos como la creación audiovisual, el entretenimiento o la educación, su potencial para el abuso plantea serios desafíos éticos y legales (Paris y Donovan, 2019). En ese sentido, Bayer et al. (2021) destacan que las consecuencias negativas de los *deepfakes* guardan semejanza con otras formas de manipulación y propaganda, pero su capacidad para crear falsificaciones extremadamente convincentes aumenta su potencial disruptivo. La popularización y viralidad de los *deepfakes* favorece su recepción, con independencia de su veracidad, agudizando este fenómeno (Vaccari y Chadwick, 2020). Esto plantea cuestiones fundamentales sobre el consentimiento y el uso indebido de contenido personal, además de la rápida difusión de material manipulado que puede vulnerar el derecho a la propia imagen y causar importantes daños reputacionales. El uso malintencionado de los *deepfakes* como herramienta de desinformación y propaganda manipulativa tiene el potencial de socavar profundamente la confianza en la información y desestabilizar el tejido social de manera particularmente grave. La verosimilitud de los *deepfakes* choca con la percepción habitual de la imagen como prueba irrefutable, esto existe porque ha sido fotografiado, generando inquietud en relación a su uso en el marco de desinformación al tiempo que transforman la verdad en una realidad mutable (Schick, 2020).

1.2. Objetivos

La investigación parte del objetivo general de analizar las repercusiones éticas y legales derivadas de la creación y de la difusión de *deepfakes,* enfocado específicamente en la problemática de la falsificación de identidades*.* Este análisis se centrará en tres objetivos específicos:

1. Identificar cómo los *deepfakes* pueden vulnerar la imagen, dignidad y privacidad de las personas generando daños reputacionales y psicológicos para las víctimas, perpetuando estereotipos y fomentando la desinformación.
2. Revisar las políticas y los mecanismos de autorregulación que implementan las plataformas de redes sociales con el fin de detectar y limitar la distribución de *deepfakes*, valorando su eficacia y proponiendo mejoras.
3. Proponer recomendaciones éticas y normativas que fomenten la transparencia, la responsabilidad y la protección de los derechos fundamentales, a través de la colaboración de la industria, la academia y los entes reguladores.

1.3. Justificación

La novedad y originalidad de esta investigación radica en el abordaje sistemático de las implicaciones éticas y legales de los *deepfakes* dentro del marco de la cultura visual digital y las redes sociales. Aunque existen estudios específicos sobre manipulación de imágenes y desinformación, persiste la necesidad de integrar la perspectiva de los derechos personales (imagen, privacidad) con la urgencia de salvaguardar la confianza social en la información visual. Vaccari y Chadwick (2020) apuntan que los *deepfakes* son una herramienta poderosa para la desinformación política, lo que genera la necesidad de desarrollar estrategias regulatorias efectivas. Además, Ajder et al. (2019) subrayan la creciente explotación de esta tecnología para fines ilícitos, destacando la urgencia de una acción legislativa coordinada. "La regulación de los *deepfakes* debe equilibrar la protección de los derechos fundamentales con la promoción de la innovación tecnológica" (Chesney y Citron, 2019, p. 1795).

En un contexto donde las tecnologías de inteligencia artificial evolucionan rápidamente y las regulaciones todavía son incipientes, este trabajo contribuye a la comprensión multidisciplinar del fenómeno y brinda recomendaciones tanto para la legislación como para la autorregulación de las plataformas digitales.

2. Metodología

El objeto formal de esta investigación es el fenómeno de los *deepfakes*, entendido como la creación de contenidos audiovisuales hiperrealistas generados mediante redes neuronales profundas (GANs), y la perspectiva adoptada considera las consecuencias éticas y legales de su implementación. Se estudia la intersección entre la tecnología que permite la manipulación de la imagen y las implicaciones que dicha manipulación entraña para la cultura visual digital contemporánea.

En este sentido, esta investigación se desarrolla siguiendo un enfoque cualitativo, el cual se fundamenta, por una parte, en una revisión sistemática de la literatura existente (Godinho Bilro y Correia Loureiro, 2020; Moher et al., 2009), cuyo objetivo es reunir datos relevantes que permitan identificar y conceptualizar las principales corrientes de investigación relacionadas con el tema estudiado (Molina Montoya, 2005; Vargas y Calvo, 1987). De este modo, se analiza una selección de ensayos que exploran los *deepfakes* como fenómenos situados en los límites de nuestro sistema de valores culturales y la tecnología.

Por otra parte, el análisis de los *deepfakes* en su contexto real se fundamenta en el enfoque del estudio de casos (Castro Monge, 2010; Yin, 1981) que buscan ilustrar el alcance de sus efectos, resultando fundamental para comprender sus implicaciones más profundas y amplias, sus significados y posibilidades (Gosse y Burkell, 2020; Paris y Donovan, 2019) profundizando en la generalización teórica (Berenguel Fernández y Fernández Gómez, 2018; Rodríguez et al., 1996).

Asimismo, nuestra comprensión de los *deepfakes* se profundizará a medida que exploremos sus continuidades históricas y puntos de ruptura con prácticas más antiguas de manipulación de medios, mediación tecnológica, fragmentación y mercantilización de imágenes humanas (Bode et al., 2021)´

3. Análisis

La recogida de datos para esta investigación se llevó a cabo siguiendo los criterios de la revisión sistemática de la literatura, recopilando información publicada entre 2018 y 2024 sobre la temática de los *deepfakes* y la cultura visual digital. Paralelamente, se seleccionaron casos de estudio bajo un criterio de relevancia mediática y debates legales y éticos suscitados, a fin de ilustrar la diversidad de usos y la magnitud de los riesgos implicados. Una vez procesada la información obtenida en la revisión bibliográfica y la correspondiente a los casos de estudio, se siguió un procedimiento de categorización de los temas emergentes (derechos de imagen, privacidad, plataformas digitales, desinformación, sesgos y narrativas discriminatorias), relacionando los hallazgos con las teorías existentes sobre la cultura visual digital y la ética de la IA. Esta triangulación entre la literatura, los datos empíricos y las propuestas legislativas facilitó la identificación de líneas de actuación recomendadas.

3.1. Abordaje del concepto de deepfake

El término *deepfake* se compone de la contracción de "deep learning" (aprendizaje profundo) y "faked imagery" (imágenes falsificadas), lo que subraya la conexión entre la tecnología de redes neuronales y la manipulación audiovisual intencionada del discurso. El origen del término se remonta a finales de 2017, cuando un usuario de Reddit llamado *Deepfakes* publicó videos pornográficos en los que se sustituían los rostros de actrices famosas por los de otras mujeres. Desde el punto de vista técnico, estos primeros *deepfakes* exigían una gran cantidad de datos y un procesamiento gráfico intensivo. Pero en enero de 2018, la aparición de *FakeApp* simplificó el proceso, haciéndolo accesible incluso para personas sin conocimientos técnicos avanzados.

Aunque el término *deepfake* se aplica a una amplia gama de formatos, suele aludir concretamente a fotografías y videos creados mediante redes neuronales profundas. En contrapartida, el concepto *cheap fake* (Paris & Donovan, 2019) se refiere a la alteración de imágenes con herramientas de edición convencionales (p. ej., *Adobe Premiere, Final Cut*), sin un alto nivel de sofisticación tecnológica.

Hoy en día la mayoría de modelos comerciales de inteligencia artificial generativa trabajan bajo estrictos filtros de control, para evitar el *deepfake*, es el caso de *Midjorney, DallE 3, Leonardo* o *Adobe Firefly,* algunas de IAs de imagen más usadas, que no permiten que las creaciones tengan parecido exacto de un rostro real dado. No obstante, existen otros portales de IA que trabajan usando modelos *FLUX* (Black Forest Labs) en sus distintas versiones, como puede ser *Grok,* un modelo que pertenece al magnate Elon Musk y que ha sido abierto de manera gratuita para todos los usuarios de la plataforma *X,* unos meses antes de las elecciones presidenciales americanas*.* Ya en modelo por suscripción está disponible la IA *Freepik* que trabaja con el modelo *FLUX pro*, el más potente de todos ellos. Las IAs generativas que trabajan usando este modelo son capaces de alcanzar niveles de fidelidad al referente original que son difícilmente detectables, lo que complica la tarea de discernir su veracidad. La mayoría de los *deepfake*s se generan en este tipo de plataformas, pero sobre todo a través de modelos locales y refinadores de IA, como *ConfyUI* y otros, bajo el motor *FLUX* para el tratamiento de los rostros, ya sea en imagen fija o creaciones audiovisuales. Dado que los motores de IA pueden descargarse de manera gratuita, los propios creadores instalan estos modelos locales en ordenadores personales de unos miles de euros y son libres para generar todo tipo de contenidos sin filtro que luego suben a la red.

3.2. Repercusiones legales y éticas de los deepfakes

Los *deepfakes* han experimentado un avance vertiginoso, generando un intenso debate tanto en el ámbito legal como en el ético, debido a su capacidad de producir contenido audiovisual extremadamente realista, pero completamente ficticio (Diakopoulos y Johnson, 2021). Esta tecnología suscita inquietudes sobre la posible vulneración de derechos fundamentales, como la privacidad y la vida personal (Rizzica, 2021), y entra en tensión con la libertad de creación artística y científica al manipular, sin consentimiento, la imagen o la voz de terceros. Su potencial para generar videos altamente verosímiles incrementa el riesgo de extorsión y difamación (Greenough, 2022), perpetúa sesgos (Mukta, Mustak y Naitali, 2023; Mustak et al., 2023), refuerza estereotipos negativos y tiene el potencial de socavar la estabilidad política (Pantserev, 2020).

No obstante, la mera elaboración de *deepfakes* no constituye, en sí misma, un delito: el elemento central para determinar la ilicitud radica en la intencionalidad y en la posibilidad de ocasionar un perjuicio concreto (Busacca y Monaca, 2023; Kirchengast, 2020; Montasari, 2024). En este sentido, algunos autores defienden la necesidad de desarrollar una inteligencia artificial "consciente" (Ng & Leung, 2020) y subrayan el papel de la gestión masiva de datos y la interoperabilidad para reforzar la protección de la privacidad (Nikolakopoulos et al., 2023). Mientras tanto, la investigación científica sigue perfeccionando métodos de detección basados en el incremento de la robustez de los detectores (Lu y Ebrahimi, 2024) y en la identificación de discrepancias posturales (Yang et al., 2019). Asimismo, desde la perspectiva periodística y de las redes neuronales, se exploran estrategias para diferenciar rostros auténticos de los simulados (Haseena et al., 2023; Shilpa et al., 2023; Sohrawardi et al., 2020). Estas prácticas, que exigen tiempo y recursos para alcanzar mayores niveles de sofisticación, generan distorsiones epistémicas y debilitan la confianza en la veracidad de las imágenes (Liz-López, 2023; Matthews, 2023).

Ante el incremento de tales riesgos, múltiples instituciones han reaccionado. El Parlamento Europeo (2021) advirtió sobre las amenazas que presentan los deepfakes en cuanto a difamación, intimidación, fraude y manipulación electoral, y señaló que entre el 90 y el 95 por ciento de estos contenidos están relacionados con la pornografía. Por su parte, la Unión Europea ha emprendido diversos mecanismos: el Reglamento (UE) 2016/679 (GDPR) establece la importancia del consentimiento expreso para el uso de datos personales y el derecho al olvido, mientras que la Ley de Servicios Digitales (Reglamento (UE) 2022/2065) impone a las plataformas la responsabilidad de salvaguardar a los usuarios y eliminar los contenidos ilegales. Adicionalmente, el Reglamento (UE) 2022/1925 relativo a los mercados digitales y el propuesto Reglamento de Inteligencia Artificial de 2021 reflejan la creciente preocupación por la generación de información falsa con fines políticos o fraudulentos.

En el plano estatal, España publicó la Proposición de Ley Orgánica para la regulación de las simulaciones de imágenes y voces de personas generadas mediante inteligencia artificial (Congreso de los Diputados, 2023), en consonancia con el informe europeo citado. Esta iniciativa busca equilibrar el derecho a la libertad de expresión con la salvaguarda del honor, la intimidad y la propia imagen. A tal efecto, se plantean reformas legales en la Ley 13/2022, de 7 de julio, General de Comunicación Audiovisual, y en la Ley Orgánica 1/1982, de 5 de mayo, con el fin de tipificar la difusión de *deepfakes* sin consentimiento expreso como infracción muy grave o como intromisión ilegítima, si no se advierte de forma inequívoca su carácter artificial. Del mismo modo, se propone reformar el Código Penal para calificar como delito de injuria la recreación de la voz o imagen de una persona mediante *deepfakes* con la intención de dañar su honor, agravándose la pena cuando la difusión se realice en redes sociales. Finalmente, se contempla la introducción de una medida cautelar en la Ley de Enjuiciamiento Civil que autorice la retirada inmediata de dichos contenidos a instancias de la persona afectada (Congreso de los Diputados, 2023).

Este panorama subraya la necesidad de un enfoque legislativo dinámico y continuo (Bode et al., 2021), en el que se incorporen las consideraciones éticas y la experiencia de los usuarios (Li & Wan, 2023), al tiempo que los sistemas jurídicos se adaptan para contrarrestar tanto el fraude como la manipulación de pruebas digitales (Montasari, 2024; Mekkawi, 2023; Mahashreshty, 2023). La posibilidad de suplantar identidades con gran verosimilitud incrementa la probabilidad de humillación y afecta la construcción de la identidad personal (Ayers, 2021). Incluso los usos recreativos o "ligeros" de la suplantación pueden perpetuar roles sociales excluyentes. Por consiguiente, la comunidad académica y los entornos regulatorios abogan por fortalecer los instrumentos de responsabilidad ante el uso malicioso de los *deepfakes*, al tiempo que se continúan investigando y desarrollando nuevas formas de detección y control de los impactos sociales.

3.3. Protección de la privacidad de los usuarios en relación con las políticas de las plataformas

La protección de la privacidad de los usuarios en entornos digitales que albergan contenido manipulado —en particular, *deepfakes*— se ha convertido en un elemento central de los debates sobre la ética en la cultura visual digital. Siguiendo a Bayer et al. (2021), la proliferación de información falsa o engañosa no solo distorsiona el funcionamiento del Estado de derecho y los procesos democráticos, sino que también vulnera la esfera personal de los individuos al exponer su imagen y datos sin consentimiento.

De este modo, la manipulación audiovisual pone en riesgo la integridad de la identidad personal, al facilitar que rostros y voces sean reutilizados con fines potencialmente ilícitos o sin autorización. Desde esta perspectiva, la detección y la limitación de la distribución de contenido manipulado se han tornado prioridades esenciales. Bode et al. (2021), en *The digital face and deepfakes on screen,* subraya la necesidad de crear políticas específicas que garanticen la salvaguarda de la privacidad de los usuarios. Además, resalta la importancia de colaborar con expertos en inteligencia artificial para desarrollar y actualizar de forma constante las herramientas de identificación de *deepfakes*, evitando su rápida obsolescencia ante la evolución de la propia tecnología. Dicho enfoque colaborativo —que involucra a corporaciones tecnológicas, investigadores independientes y organismos reguladores— resulta crítico para mitigar la propagación de contenido manipulado. Sin embargo, la eficacia de estas políticas presenta diversos desafíos. De acuerdo con Bode et al. (2021), varias plataformas han implementado algoritmos de IA para reconocer y eliminar contenidos manipulados; sin embargo, la velocidad con la que emergen y se perfeccionan las técnicas de generación de *deepfakes* dificulta la supresión efectiva antes de que causen un impacto significativo. A su vez, Ayers (2021) pone el acento en la baja transparencia de los procesos de moderación de muchas plataformas, lo que obstaculiza la verificación independiente de la eficacia de los mecanismos de detección y eliminación. Esta falta de claridad debilita la confianza tanto de la comunidad académica como de los propios usuarios, dificultando una evaluación objetiva de los resultados obtenidos. En definitiva, la protección de la privacidad en el contexto de los *deepfakes* exige un enfoque integral que combine el diseño de políticas sólidas, la colaboración sistemática con expertos en inteligencia artificial y una mayor transparencia de las plataformas. En tanto la tecnología de generación de contenidos manipulados continúe su acelerada evolución, las estrategias orientadas a salvaguardar la identidad personal y el derecho a la propia imagen deberán renovarse y fortalecerse continuamente, atendiendo tanto las dimensiones éticas como legales que sostienen la cultura visual digital (Ayers, 2021; Bayer et al., 2021; Bode et al., 2021).

A continuación, recogemos las medidas adoptadas por dos de las plataformas más importantes, Meta y X con objeto ilustrar la protección de la privacidad de sus usuarios así como su evolución. Meta (anteriormente Facebook) estableció una política con la que buscaba combatir la alteración de la identidad personal a través de IA y prohibir, en esencia, los usos malintencionados de *deepfakes* (Meta Platforms, Inc., 2024). Esta iniciativa incluía herramientas de detección automatizadas, permitiendo, sin embargo, la publicación de contenido manipulado con fines satíricos o educativos, siempre que se etiquetara de forma clara. A finales de 2023, Marc Zuckerberg anunció la eliminación de la verificación externa por parte de terceros, alegando que el fact checking independiente equivalía a una forma de censura (Zuckerberg, 2023). A pesar de ello, en España siguen trabajando organizaciones como la Fundación Maldita y Newtral para verificar publicaciones (García, 2023). Por su parte, X (antes Twitter) también contemplaba políticas de etiquetado para contenidos potencialmente manipulados, alertando a los usuarios sobre la naturaleza modificada del material antes de su difusión general (Twitter, 2023). No obstante, e igualmente en 2023 relajó ciertas restricciones y redujo la colaboración con entidades de verificación, generando un aumento de la desinformación en la plataforma. En este escenario, la Comisión Europea abrió una investigación formal contra X por supuestas infracciones relacionadas con la moderación de contenido y la transparencia, en el marco de la Ley de Servicios Digitales (DSA), que exige mayores esfuerzos para eliminar contenidos ilegales y preservar la privacidad y la información veraz (Pérez, 2023).

3.4. Casos de estudio de uso malintencionado de deepfakes

Recogemos seguidamente algunos casos significativos de *deepfakes* malintencionados en España, Europa y Estados Unidos, reportados en la prensa y por organizaciones dedicadas a la verificación de información en los dos últimos años. Cada situación presenta características específicas, pero en todas ellas las víctimas experimentan repercusiones en su esfera personal, social y profesional.

1. *Deepfakes* de contenido sexual de menores en Almendralejo (Badajoz). En el verano de 2023, se reveló que al menos 20 menores habían sido víctimas de compañeros de clase que utilizaban IA para generar imágenes de contenido sexual falso a partir de sus fotografías cotidianas. El suceso, aún en investigación, abrió el debate sobre la carencia de legislación específica y los vacíos legales en España respecto al uso de *deepfakes*. Se han planteado peticiones de reforma del Código Penal para abordar la suplantación de identidad y la generación de contenidos falsos. Este incidente

pone en evidencia la vulnerabilidad de los adolescentes frente a este tipo de manipulación digital, que no solo atenta contra su dignidad, sino que también genera consecuencias psicológicas profundas, como vergüenza, culpa y estigmatización dentro de su entorno escolar. El temor a ser rechazados por sus compañeros y a que circulen nuevas versiones de los *deepfakes* agrava su ansiedad, mientras los sistemas legales aún no ofrecen medidas suficientemente claras o inmediatas para su protección.

Imagen 1. Captura de pantalla del artículo El caso de los falsos desnudos de menores de Almendralejo.


elDiario.es

Hazte socio/a Inicia sesión

Extremadura

Política Sociedad Economía Educación Igualdad en derechos Cultura Turismo Energías renovables Territo

El caso de los falsos desnudos de menores de Almendralejo generados por inteligencia artificial llegará a Bruselas

Miriam Al Adib, madre de una de las víctimas, intervendrá en el Parlamento Europeo para intentar que la violencia sexual cometida a través de la IA forme parte de la agenda política

— La Fiscalía investiga a 26 menores en el caso de los falsos desnudos de niñas en Almendralejo

Fuente: elDiario.es (2023, 8 de noviembre). Recuperado de El caso de los falsos desnudos de menores de Almendralejo generados por inteligencia artificial llegará a Bruselas

2. *Deepfakes* pornográficos no consentidos de influencers y figuras públicas. En Europa se han reportado casos de *deepfakes* pornográficos no consentidos protagonizados por celebridades e influencers. En cuanto a los daños reputacionales y psicológicos, al difundirse grabaciones falsas con contenido sexual, las víctimas padecen humillación, pérdida de credibilidad y perjuicio laboral. Este tipo de violencia digital de género y la denominada "revenge porn" agrava la desigualdad de género y repercute con mayor fuerza en la esfera emocional y social de las mujeres, sometidas al escrutinio público en un entorno que a menudo las responsabiliza injustamente de estas agresiones. Se ha exigido mayor protección legal y herramientas para salvaguardar a las víctimas de estos ataques.

Imagen 2. Captura de pantalla del artículo *De Emma Watson a Alexandria Ocasio-Cortez: cómo la IA se está usando para crear imágenes y vídeos porno o que sexualizan a mujeres famosas sin su consentimiento*.

*F*uente: Maldita Tecnología. (2025, 22 de enero). Recuperado de https://maldita.es/malditatecnologia/20250122/ia-imagenes-porno-mujeres-famosas/

3. *Deepfakes* de figuras públicas. La suplantación de identidad de figuras públicas en Europa ha sido igualmente alarmante, tal como refleja el caso del alcalde de Madrid, quien llegó a mantener una conversación virtual con un *deepfake* del alcalde de Kiev, Vitali Klitschko. Sucesos similares afectaron también a la alcaldesa de Berlín y al alcalde de Viena. Estas acciones buscan desacreditar a personalidades públicas y difundir desinformación. Así mismo, el temor a ser víctimas de ataques similares produce un alto grado de estrés exponiendo a estas figuras a cuestionamientos constantes sobre la autenticidad de sus apariciones, lo que dificulta su actividad pública y desgasta

su reputación e incluso la credibilidad de las instituciones que representan. La Unión Europea ha reforzado la cooperación entre instituciones y discute la necesidad de tipificar la creación y difusión de *deepfakes* maliciosos como delito específico.

Imagen 3. Captura de pantalla del artículo *Almeida, víctima de un impostor que se hizo pasar por el alcalde de Kiev.*

Fuente: La Razón (2022, 25 de junio). Recuperado de Almeida, víctima de un impostor que se hizo pasar por el alcalde de Kiev

4. *Deepfakes* en campañas políticas y desinformación electoral en EE.UU. En campañas políticas de Estados Unidos han proliferado los *deepfakes* con fines de desinformación, manipulando fragmentos de discursos o atribuyendo mensajes extremistas a candidatos. En ese contexto, desde 2023, organizaciones de verificación de datos alertaron sobre el aumento de vídeos *deepfake* usados en contextos políticos. Este uso malintencionado de la tecnología erosiona la confianza en las instituciones democráticas y genera una atmósfera de escepticismo y confrontación política. Los candidatos y sus equipos enfrentan estrés adicional al verse obligados a desmentir falsedades de manera constante, lo que reduce su capacidad de concentrarse en la comunicación política legítima. La sociedad, a su vez, se ve arrastrada a un clima de polarización y sospecha que afecta la estabilidad mental colectiva. Ante tales hechos, redes sociales como Twitter/X y Facebook/Meta reforzaron políticas de moderación y etiquetado de contenido manipulado, aunque la eficacia de estas medidas sigue en entredicho.

Imagen 4. Captura de pantalla de la noticia *Las imágenes falsas creadas con IA para intentar atraer el apoyo de los votantes negros en EE.UU.*

Fuente: BBC News Mundo. (2023, 1 de abril). Recuperado de https://www.bbc.com/mundo/articles/c3g4l5xgvryo

4. Conclusiones y discusión

Los *deepfakes* constituyen un fenómeno de alto impacto en la cultura visual digital contemporánea, con la capacidad de dañar la imagen y reputación, vulnerar la privacidad y debilitar la credibilidad de la información audiovisual. La investigación llevada a cabo muestra cómo la facilidad de crear *deepfakes* hiperrealistas expone tanto a figuras públicas como a personas anónimas a estas manipulaciones, cuyo rostro y voz pueden ser manipulados sin consentimiento, comportando daños psicológicos para ellas. Por otra parte, la creciente accesibilidad de las *deepfakes* combinada con el entorno digital hiperconectado y la viralidad de las redes sociales acelera su difusión masiva, dificultando la eliminación de contenido falsificado. En ese aspecto, los *deepfakes* amplían la efectividad de las campañas de desprestigio o engaño político, socavando la confianza ciudadana en la veracidad de los materiales audiovisuales y generando un escepticismo generalizado entre los ciudadanos. Además, se constata que la manipulación audiovisual puede dirigir ataques a colectivos específicos, perpetuando estereotipos y fomentando la violencia simbólica en redes. Respecto a las plataformas, a pesar de los esfuerzos en el desarrollo de métodos automáticos de detección, la rápida evolución de los *deepfakes* amplía la brecha entre creadores y las plataformas. En ese aspecto, aunque se identifican esfuerzos crecientes en el desarrollo de métodos automáticos de detección y la propuesta de marcos regulatorios específicos, su aplicación práctica y su eficacia a gran escala requieren de mayor desarrollo y armonización. Este vacío legal, unido al perfeccionamiento constante de las herramientas tecnológicas, dificulta la respuesta de las plataformas digitales y deja a las víctimas sin protección adecuada. Además, los *deepfakes* no solo amenazan la privacidad individual, sino que también tienen implicaciones políticas y sociales significativas, desde la interferencia en campañas electorales hasta la perpetuación de discursos de odio. Por todo ello, se requiere un enfoque integral donde confluyan la legislación, la ética, la responsabilidad de las plataformas digitales y la formación de la ciudadanía en materia de verificación y consumo crítico de contenidos audiovisuales.

En línea con autores como Chesney y Citron (2019) se observa cómo los *deepfakes* contribuyen a una "nueva guerra de la desinformación" que se abre con la posibilidad de manipulación audiovisual de forma prácticamente indetectable. La constancia de casos de difamación y abusos corrobora las reflexiones de Vaccari & Chadwick (2020) sobre la erosión de la confianza pública en la veracidad de la información, y subraya el riesgo de un escepticismo sistémico frente a todo contenido visual. A su vez, Bode et al. (2021) plantean la complejidad de las plataformas digitales para enfrentar una tecnología en constante perfeccionamiento, reforzando lo que se ha observado en la limitada efectividad de las políticas de moderación.

La discusión se extiende a la dimensión legal, donde las aproximaciones de Kirkengast (2020) y Meskys et al. (2020) evidencian la urgencia de concretar leyes que aborden la manipulación audiovisual profunda, superando la aplicación residual de delitos tradicionales. Li y Wan (2023) enfatizan también la necesidad de contemplar la experiencia del usuario y la ética en la creación de estos contenidos, pues, como hemos visto, no todo *deepfake* se realiza con finalidades nocivas.

En la esfera más amplia de la cultura visual digital, estas conclusiones concuerdan con la tesis de Bañuelos (2022) sobre la rápida transformación del ecosistema mediático y de Ajder et al. (2019) respecto a la capacidad de los *deepfakes* para generar falsificaciones muy verosímiles. En línea con lo expuesto por Brown & Fleming (2020) y Maddocks (2020), el uso malintencionado de esta tecnología refuerza la violencia de género y otras modalidades de hostigamiento, evidenciando cómo la IA puede amplificar dinámicas de poder preexistentes y perpetrar sesgos de todo tipo.

Por último, se constata la pertinencia de la perspectiva de "deterioro de la verdad" (Fletcher, 2018) en la era posfactual: la mera posibilidad de que un contenido sea un *deepfake* extiende la duda hacia cualquier material audiovisual. Este escenario invita a una reflexión sobre la responsabilidad compartida de gobiernos, desarrolladores de IA, plataformas digitales y usuarios para desplegar técnicas de detección cada vez más precisas (Rössler et al., 2019; Tolosana et al., 2020), al tiempo que se impulsa una alfabetización digital que capacite a la sociedad para una lectura crítica de las imágenes y los vídeos que consume.

En síntesis, tras la revisión de la bibliografía existente y el análisis de los casos de estudio, podemos concluir que la conjunción de resultados empíricos y debates teóricos confirma la necesidad de un enfoque integral para frenar el uso malintencionado de los *deepfakes* en la cultura visual digital, posicionando el foco en tres vías de actuación prioritarias:

En primer lugar, se requiere una legislación adecuada que no solo prohíba la creación y difusión de *deepfakes* malintencionados, sino que también se adapte de forma continua a los avances tecnológicos. Esto incluye imponer multas significativas a plataformas que permitan la proliferación de este tipo de contenidos sin ofrecer herramientas eficaces de detección o etiquetado. Un ejemplo de ello es la regulación europea, que con la Ley de Servicios Digitales (DSA) obliga a las plataformas de gran tamaño (Very Large Online Platforms) a identificar, analizar y mitigar riesgos sistémicos como la desinformación, el discurso de odio y los contenidos ilegales. Esto incluye trabajar con organizaciones de fact-checking (verificación de hechos) para detectar y limitar la propagación de información falsa. Como ejemplo añadido cabría recordar la legislación China, que obliga a que los contenidos generados por IA incluyan marcas de agua visibles o metadatos identificativos, penalizando a las plataformas que incumplen estos requisitos (Cyberspace Administration, 2024). De igual modo, iniciativas como la Ley AB 730 de California, que restringe los *deepfakes* en contextos electorales, demuestran el potencial de las normativas específicas para mitigar riesgos en áreas sensibles (AB 730, 2019). La creación de un marco normativo global que responsabilice a las plataformas digitales y garantice transparencia es una necesidad urgente.

En segundo lugar, la industria tecnológica y los agentes externos deben asumir un rol activo mediante la implementación de estándares de transparencia y la creación de herramientas avanzadas de detección (Sanchez et al., 2024). Protocolos como los desarrollados por Adobe en su Content Authenticity Initiative permiten rastrear y verificar el origen de contenidos generados por IA mediante marcas de agua digitales. Además, iniciativas como la herramienta Video Authenticator de Microsoft, diseñada para analizar vídeos en busca de manipulaciones, muestran cómo las empresas tecnológicas pueden contribuir al control de este fenómeno (Microsoft, 2020). La colaboración entre grandes actores tecnológicos, como en el caso del consorcio Partnership on AI, establece un modelo de cooperación que debería ampliarse para incluir un repositorio global de soluciones accesible para gobiernos y usuarios (Partnership on AI, 2020).

Por último, es imprescindible aumentar la educación visual y la alfabetización mediática para dotar a la ciudadanía de herramientas críticas frente a los *deepfakes*. Programas educativos como los implementados en Finlandia, que integran la alfabetización mediática en todos los niveles del sistema educativo, destacan por su eficacia en la formación de ciudadanos críticos y resilientes frente a la desinformación (European Commission, 2019). Estas iniciativas deberían ampliarse a nivel global, incorporando módulos específicos sobre verificación de contenidos, formación docente y campañas de concienciación pública, etcétera. Educar a la población para reconocer manipulaciones digitales es por tanto esencial para fortalecer la sociedad frente a la desinformación (Ajder et al. 2019).

En el ámbito de la sensibilización y la alfabetización mediática en torno a los *deepfakes,* uno de los primeros ejemplos notables fue la producción de Jordan Peele junto con BuzzFeed (Peele & BuzzFeed, 2018), en la cual se mostraba al expresidente Barack Obama emitiendo declaraciones falsas. Mediante el uso de inteligencia artificial para sincronizar la voz y los gestos de Peele con la imagen de Obama, este proyecto buscó evidenciar con fines educativos la facilidad de manipular contenido audiovisual.

Imagen 5. Captura de pantalla del video en el que Jordan Peele *You Won't Believe What Obama Says in This Video!*

Fuente: Peele & BuzzFeed, (2018).

De forma similar, la exposición *A chupar del bote* de Joan Fontcuberta (2017) —atribuyéndosela a un reportero ficticio— mostró documentos y retratos manipulados para subrayar la urgencia de un consumo crítico de imágenes (Fontcuberta y Fernández, 2017).

Imagen 6. Imagen de la exposición *A chupar del bote* de Ximo Berenguer (Joan Fontcuberta).

Fuente: Fontcuberta, J. (2017). Exposición en PhotoEspaña 2017.

Esta preocupación también fue recogida en la muestra *Fake News. La fábrica de mentiras* (2023-2024), expuesta en el Espacio Fundación Telefónica, que abordó la desinformación desde una perspectiva histórica y propuso un decálogo para combatirla (Fundación Telefónica Movistar, 2024).

Imagen 7. Vista de la exposición *Fake News. El valor de la información* en la Fundación Telefónica Movistar, Buenos Aires, 2024.

Fuente: Fundación Telefónica Movistar. (2024).

Por último, la iniciativa *This Person Does Not Exist* (ThisPersonDoesNotExist.com, s.f.) utilizó redes neuronales generativas para crear rostros hiperrealistas de individuos inexistentes, poniendo de relieve el potencial de la IA para generar identidades ficticias y promoviendo la reflexión sobre la ética y los riesgos asociados a estas prácticas.

Imagen 8. Rostro generado por inteligencia artificial, usando IA y sin representar a ninguna persona real.

Fuente: ThisPersonDoesNotExist.com (s.f.). Recuperado de https://thispersondoesnotexist.com.

En definitiva, para abordar los desafíos que plantea la proliferación de *deepfakes* se precisa de estos tres factores, que trabajan armónicamente en codependencia. La experiencia de países como China, Estados Unidos y Finlandia demuestra que es posible implementar medidas efectivas cuando se actúa desde múltiples frentes. Sin embargo, la coordinación internacional y la colaboración entre gobiernos, industria tecnológica y ciudadanía serán determinantes para garantizar que estas iniciativas tengan un impacto real en la construcción de un ecosistema digital más seguro y transparente.

Referencias

Ajder, H., Patrini, G., Cavalli, F., & Cullen, L. (2019). The State of Deepfakes: Landscape, Threats, and Impact. Deeptrace Labs Report. Disponible en:
https://regmedia.co.uk/2019/10/08/deepfake_report.pdf

Assem. Bill 730, Ch. 493. (Cal. 2019). https://leginfo.legislature.ca.gov/faces/billTextClient.xhtml?bill_id=201920200AB730

Ayers, D. (2021). The limits of transactional identity: Whiteness and embodiment in digital facial replacement. Convergence, 27(4), 1018-1037. 10.1177/13548565211027810

Bañuelos, J. (2022). Evolución del Deepfake: campos semánticos y géneros discursivos (2017-2021). Revista ICONO 14. Revista Científica De Comunicación Y Tecnologías Emergentes, 20(1). https://doi.org/10.7195/ri14.v20i1.1773

Baudrillard, J. (1999). Sur la photographie. París: Sens & Tonka.

Bayer, J., Holznagel, D. B., Katarzyna, L., Pace, A., Szakács, J., & Uszkiewicz, E. (2021). Disinformation and propaganda: Impact on the functioning of the rule of law and democratic processes in the EU and its Member States: 2021 update. European Parliament Policy Department for External Relations. https://bit.ly/3WHc8ZM

Berenguel Fernández, J., & Fernández Gómez, J. (2018). La eficacia de la comunicación en la convergencia mediática: propuesta de metodología de estudio y aplicación de casos. Trípodos, (43), 37-56.

Berenguer, X. (2016). A chupar del bote. RM Verlag

Bode, L., Lees, D., & Golding, D. (2021). The Digital Face and Deepfakes on Screen. Convergence, 27(4), 849-854. 10.1177/13548565211034044

Brown, M., & Fleming, E. (2020). Celebrity headjobs: Or oozing squid sex with a framed-up leaky. Porn Studies, 7(4), 357-366. http://dx.doi.org/10.1080/23268743.2020

Busacca, A., & Monaca, M. A. (2023). Deepfake: Creation, Purpose, Risks. En Marino, D. y Monaca, M. A. (eds.), Innovations and Economic and Social Changes due to Artificial Intelligence: The State of the Art (pp. 55-68). Springer Nature Switzerland. 10.1007/978-3-031-334610_6

Buxó, M.J. (1999). De la investigación audiovisual. Fotografía, cine, vídeo, televisión. Barcelona: Proyecto A Ediciones.

Castro Monge, E. (2010). El estudio de casos como metodología de investigación y su importancia en la dirección y administración de empresas. Revista Nacional de Administración, 1(12), 31-54. doi: https://doi.org/10.22458/rna.v1i2.332

Chesney, R., & Citron, D. (2018). Deep Fakes: A Looming Challenge for Privacy, Democracy, and National Security. California Law Review, 107(6), 1753-1816.
DOI: https://doi.org/10.2139/ssrn.3213954

Chesney, R., & Citron, D. (2019). Deepfakes and the New Disinformation War: The Coming Age of Post-Truth Geopolitics. Foreign Affairs.
Disponible en: https://www.jstor.org/stable/26798018

Citron, D. (2014). Hate crimes in cyberspace. Harvard University Press.

Congreso de los Diputados. (2023). Proposición de Ley [BOCG, Serie B, 15-B-23-1].https://bit.ly/40CHHFn

Cyberspace Administration of China. (2024). Measures for the Identification of AI-Generated Synthetic Content [Medidas para la identificación de contenido sintético generado por IA]. http://www.cac.gov.cn/

Diakopoulos, N., & Johnson, D. (2021). Anticipating and addressing the ethical implications of deepfakes in the context of elections. New Media & Society, 23(7), 2072-2098. 10.1177/1461444820925811

European Commission. (2019). Youth policies in Finland: 2019. European Education and Culture Executive Agency.

Fernández, Á. (2017). Relatos híbridos: El papel de la narratividad en la visualización de información interactiva [Tesis doctoral, Universidad Europea]. Repositorio Abacus https://193.147.239.238/handle/11268/6981

Fletcher, R. (2018). The truth behind fake news in a post-factual age. Palgrave Studies in Communication.

Fontcuberta, J. (2017). A chupar del bote [Exhibición]. Galería Fernando Pradilla, Madrid, España.

Fontcuberta, J., y Fernández, H. (2017). A chupar del bote. Editorial RM.
Fundación Telefónica. (2023). Fake News. La fábrica de mentiras [Exposición]. Madrid, España: Espacio Fundación Telefónica. Recuperado de https://bit.ly/3EeEBQo
García, M. (2023, octubre 15). Zuckerberg elimina verificación de datos en Facebook e Instagram, modelo de Elon Musk en X. ElDiario.es. https://bit.ly/40TiUhM
Godinho Bilro, R., & Correia Loureiro, S. (2020). A consumer engagement systematic review: synthesis and research agenda. Spanish Journal of Marketing - ESIC, 24(3), 283-307.
Gosse, C., & Burkell, J. (2020). Politics and porn: how news media characterizes problems presented by deepfakes. Critical Studies in Media Communication, 37(5), 497-511.
Greenough, C. J. (2022). Make it, Fake it and Get Away with it? The Role of Toxic Masculinity and Threat Perception within Cases, Policies, and Legislation Surrounding Deep Fakes. Master Thesis, University of Auckland. https://hdl.handle.net/2292/61261
Haseena, S., Saroja, S. & Nivetha, A. (2023). TVN: Detect Deepfakes Images using Texture Variation Network. Inteligencia Artificial, 26(72), 1–14. https://doi.org/10.4114/intartif.vol26iss72pp1-14
Kirchengast, T. (2020). Deepfakes and image manipulation: Criminalisation and control. Information & Communications Technology Law, 29(3), 308-323. 10.1080/13600834.2020.1794615
Li, M., & Wan, Y. (2023). Norms or fun? The influence of ethical concerns and perceived enjoyment on the regulation of deepfake information. Internet Research, 33(5), 1750-1773. 10.1108/INTR-07-2022-0561
Liz-López, H, Keita, M., Taleb-Ahmed, A. Abdenour, H., Huertas-Tato, J., Camacho, D. (2023). Generación y detección de contenidos audiovisuales multimodales manipulados: Avances, tendencias y desafíos abiertos. Fusión de Información, pp.102-103
Lu, G., & Ebrahimi, T. (2024). A New Approach to Improve Learning-based Deepfake Detection in Realistic Conditions. arXiv preprint arXiv:2203.11807.
Maddocks, S. (2020). Deepfake Porn, Revenge Porn, and Gendered Violence. Feminist Media Studies, 20(7), 976–990.
Mahashreshty, V. S. (2023). Implications of deepfake technology on individual privacy and security (Culminating Projects in Information Assurance No. 142). St. Cloud State University. https://repository.stcloudstate.edu/msia_etds/142
Matthews, T. (2023). Deepfakes, fake barns, and knowledge from videos. Synthese, 201 (2). https://doi.org/10.1007/s11229-022-04033-x.
Mekkawi, L. (2023). The challenges of Digital Evidence usage in Deepfake Crimes Era. Journal of Law and Emerging Technologies, 3(2), 176-232.
Meskys, E., Kalpokiene, J., Jurcys, P., & Liaudanskas, A. (2020). Regulating deep fakes: legal and ethical considerations. Journal of Intellectual Property Law & Practice, 15(1), 24-31.
Meta Platforms, Inc. (2024, 5 de abril). Our approach to labeling AI-generated content and manipulated media. Meta. https://about.fb.com/news/2024/04/metas-approach-to-labeling-ai-generated-content-and-manipulated-media/
Microsoft. (2020, septiembre 2). Microsoft anuncia novedades en su programa Defending Democracy para luchar contra las campañas de desinformación. Microsoft News Center. https://news.microsoft.com/es-es/2020/09/02/microsoft-anuncia-novedades-en-su-programa-defending-democracy-para-luchar-contra-las-campanas-de-desinformacion/
Ministerio del Interior (2023). Informe de Cibercriminalidad en España 2023. (Referencia sin URL oficial.)
Moher, D., Shamseer, L., Clarke, M., Ghersi, D., Liberati, A., Petticrew, M., … PRISMA-P Group. (2009). Preferred reporting items for systematic review and meta-analysis protocols (PRISMA-P) 2015 statement. Syst Rev, 4(1).
Molina Montoya, N. P. (2005). Herramientas para investigar: ¿Qué es el estado del arte? Ciencia y Tecnología para la Salud Visual y Ocular, (4), 73-75.
Montasari, R. (2024). Responding to Deepfake Challenges in the United Kingdom: Legal and Technical Insights with Recommendations. En Montasari, R. (ed.), Cyberspace, Cyberterrorism and the International Security in the Fourth Industrial Revolution: Threats, Assessment and Responses (pp. 241-258). Springer International Publishing. 10.1007/978-3-031-50454-9_12

Mukta, N., Mustak, M., & Naitali, A. (2023). An investigation of the effectiveness of deepfake models and tools. Journal of Sensor and Actuator Networks, 12(4), 61. https://doi.org/10.3390/jsan12040061

Mustak, M., Salmien, J., Mäntymäki, M., Rahman, A. & Dwivedi, Y.K. (2023). Deepfakes: Deceptions, Mitigations and Opportunities. Journal of Business Research, 154.

Ng, A., & Leung, M. (2020). Toward a strong AI for deepfake governance: Ethical implications. (Referencia sin datos.)

Nikolakopoulos, A., Segui, M., Pellicer, A. B., Kefalogiannis, M., Gizelis, C., Marinakis, A., Nestorakis, K., & Varvarigou, T. (2023). BigDaM: Efficient Big Data Management and Interoperability Middleware for Seaports as Critical Infrastructures. Computers, 12(11), Article 11. https://doi.org/10.3390/computers12110218

Pantserev, K. A. (2020). The Malicious Use of AI-Based Deepfakes Technology as the New Threat to Psychological Security and Political Stability. En Jahankhani, H., Kendzierskyj, S., Chelvachandran, N. y Ibarra, J. (eds.), Cyber Defence in the Age of AI, Smart Societies and Augmented Humanity (pp. 37-55). Springer International Publishing. 10.1007/978-3-030-35746-7_

Paris, B., & Donovan, J. (2019). Deepfakes and Cheap Fakes: The Manipulation of Audio and Visual Evidence. Data & Society Research Institute.

Disponible en: https://datasociety.net/library/deepfakes-and-cheap-fakes/

Partnership on AI. (2020, 12 de marzo). The Deepfake Detection Challenge: Insights and recommendations for AI and media integrity. https://partnershiponai.org/wp-content/uploads/2021/07/671004_Format-Report-for-PDF_031120-1.pdf

Peele, J., & BuzzFeed. (2018). You Won't Believe What Obama Says In This Video! [Video]. YouTube. https://www.youtube.com/watch?v=cQ54GDm1eL0

Pérez, J. (2023, diciembre 18). Bruselas abre una investigación formal contra X por vulnerar las normas sobre moderación de contenidos. Cinco Días. https://bit.ly/3El4b6n

Rizzica, F. (2021). Protección del bienestar en la era de los deep fakes. (Referencia sin datos concretos.)

Rodríguez, G., Gil, J., & García, E. (1996). Metodología de la investigación cualitativa. Ediciones Aljibe.

Rössler, A., Cozzolino, D., Verdoliva, L., Riess, C., Thies, J., & Niessner, M. (2019). FaceForensics++: Learning to detect manipulated facial images. Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV).

Sánchez, M., Palella, S., & Fernández, A. (2024). Implementación de herramientas de Inteligencia Artificial en la detección de vídeos falsos y ultrafalsos (deepfakes): Caso de Radio Televisión Española (RTVE). VISUAL REVIEW. International Visual Culture Review/Revista Internacional de Cultura Visual, 16(4), 213-225.

Schick, N. (2020). Deepfakes: the coming infocalypse. Grand Central Publishing.

Shilpa, A., Varma, K., & Sur, S. (2023). Evaluating convolutional neural networks for deepfake detection. (Referencia sin datos concretos.)

Sohrawardi, S. J., Seng, S., Chintha, A., Thai, B., Hickerson, A., Ptucha, R., & Wright, M. (2020). Defaking DeepFakes: Understanding journalists' needs for DeepFake detection. In Proceedings of the Computation+ Journalism 2020 Conference (Vol. 21). Northeastern University.

Tolosana, R., Vera-Rodríguez, R., Fierrez, J., Morales, A., & Ortega-Garcia, J. (2020). Deep Fakes and beyond: A survey of face manipulation and fake detection. Information Fusion, 64, 131-148.

UNESCO (2022). Recomendación sobre la Ética de la Inteligencia Artificial. https://unesdoc.unesco.org/ark:/48223/pf0000380455

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).*1-13. https://doi.org/10.1177/2056305120903408

Vargas, J., & Calvo, V. (1987). El estado del arte en la investigación educativa: su definición y significado. Perfiles Educativos, 35, 5-15.

Wagner, T. & Blewer, A. (2019). "The Word Real Is No Longer Real": Deepfakes, Gender, and the Challenges of AI-Altered Video. Open Information Science, 3(1), 32-46. https://doi.org/10.1515/opis-2019-0003

Wang, P. (2019). This person does not exist [Portal web]. Recuperado de https://thispersondoesnotexist.com

Westerlund, M. (2019). The Emergence of Deepfake Technology: A Review. Technology Innovation Management Review, 9(11), 39-52.
DOI: https://doi.org/10.22215/timreview/1282
X. Yang, Y. Li, & S. Lyu (2019). Exposing deep fakes using inconsistent head poses. ICASSP 2019 - 2019 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), 8261–8265.
Yin, R. (1981). The case study crisis: Some answers. Administrative Science Quarterly, 26(1), 58-65
Zuckerberg, M. (2023, octubre 20). Actualización sobre las nuevas funcionalidades de Facebook [Video]. Facebook. https://www.facebook.com/zuck/videos/1525382954801931