



## Los *deepfakes* y su incidencia en el derecho probatorio

### The deepfakes and their impact on evidence law

#### RESUMEN

El *objetivo* de este artículo es determinar el impacto de las tecnologías *deepfake* en el derecho probatorio, debido a que la inteligencia artificial generativa no solo provoca un estado de indefensión para bienes jurídicos como la democracia, sino que también se configura como una problemática disruptiva para la verdad procesal. Asimismo, a lo largo de la presente investigación se empleó el Método de Revisión Integrativa de Literatura mediante el cual se recolectó y filtró información vinculada a la naturaleza e impacto de los ultrafalsos. Finalmente, se llegó a demostrar que en el *common law*, se debería considerar la posibilidad de erradicar el uso del Principio del Testigo Silencioso frente a las pruebas digitales en consideración de la proliferación de los *deepfakes*; y, por otro lado, en el *civil law*, se pudo evidenciar que la necesidad de utilizar inteligencia artificial para la detección de *deepfakes* atentaría contra dos criterios Daubert, ya que la técnica carece de pruebas frente a errores y le hace falta revisión científica.

**Palabras clave:** *deepfakes*; derecho probatorio; derecho procesal; noticias falsas; ultrafalsos.

#### ABSTRACT

The *objective* of this study is to determine the impact of deepfake technologies on the law of evidence, since generative artificial intelligence not only causes a state of defenselessness for legal assets such as de-

---

<sup>1</sup> Bachiller en Derecho y Administración y Negocios Internacionales por la Universidad La Salle de Arequipa – Perú. Egresado de la Maestría en Tributación y Política Fiscal por la Universidad Nacional de San Agustín. Egresado de la Maestría en Bioética y Biojurídica por la Universidad Santo Toribio de Mogrovejo. Maestrando en el MBA en Gerencia de Proyectos de Inversión por la Universidad Nacional de San Agustín; ORCID: <https://orcid.org/0000-0002-0422-0954>; correo electrónico: [coaquira@ulasalle.edu.pe](mailto:coaquira@ulasalle.edu.pe)

# LOS “DEEPFAKES” Y SU INCIDENCIA EN EL

DERECHO PROBATORIO

mocracy, but also constitutes a disruptive problem for procedural truth. Likewise, throughout this research, the Integrated Literature Review (ILR) Method was used, through which information related to the nature and impact of deepfakes was collected and filtered. Finally, it was shown that in common law, the possibility of eradicating the use of the Silent Witness Principle should be considered in relation to digital evidence in consideration of the proliferation of deepfakes; and, on the other hand, in civil law, it was possible to show that the need to use artificial intelligence for the detection of deepfakes would violate two Daubert criteria, since the technique lacks evidence against errors and lacks scientific review.

**Keywords:** deepfakes; evidence law; fake news; generative artificial intelligence; procedural law.

## INTRODUCCIÓN

**N**o existe área sobre la que se desenvuelva el ser humano que no se haya visto afectada por la disrupción de la inteligencia artificial (IA). Como consecuencia de su desarrollo exponencial, se han desprendido un cúmulo de tecnologías que han planteado una serie de beneficios y problemáticas para la sociedad, desde el desplazamiento de la mano del hombre por la automatización inteligente hasta la generación de videos, fotos o audios alterados por la IA denominados *deepfakes* (Mustak et ál., 2023, pp. 6-11).

Si bien hasta el día de hoy los *deepfakes* han sido empleados en su mayoría para la generación de pornografía no consensuada o memes, su potencial uso aún no ha dividido límites, ya que dentro de las posibles utilidades beneficiosas que se le podrían dar se encuentran el apoyo a la industria cinematográfica para la gestión de escenas con actores fallecidos, la remodelación tanto del comercio electrónico como del ámbito publicitario, el aumento de telepresencia en chats virtuales, entre otros.

Por otro lado, esta tecnología no solo ha generado desinformación a grandes escalas por causa del alcance desmedido de las redes sociales, sino que ha promovido un apocalipsis de la información o una apatía ante la realidad, lo cual es más alarmante. En otras palabras, los medios sintéticos dentro de los que destacan los *deepfakes* han concretado la era de la pos-verdad, la cual estaría caracterizada por la primacía de lo sentimental ante lo objetivo (Bañuelos, 2020, pp. 53-54). Si bien la desinformación puede generar un estado de indefensión para empresas, sistemas políticos o cualquier elemento de la sociedad, estas nuevas tecnologías se configuran como una cuestión alarmante para el derecho procesal, sobre todo para el derecho probatorio, ya que si bien el objeto de la prueba sería generar certeza en un juez, la posibilidad de que un *deepfake* ingrese a un proceso haría cada vez más inviable dicha finalidad. Cabe resaltar que un determinado sector afirma que la posibilidad de detectar *deepfakes*, se hace cada vez menor, debido a los avances de los sistemas inteligentes.

En ese sentido, el *objetivo* de la presente investigación es analizar el impacto de los *deepfakes* en el ámbito de la prueba. Siendo así, se compararán las políticas y regulaciones en torno a los *deepfakes* que han sido promovidas a escala internacional. Así también se evaluarán las potenciales pro-

blemáticas provenientes de la actuación de *deepfakes* como medios probatorios.

Así pues, el estudio mantiene la siguiente estructura: primero se desarrolla la metodología de Revisión Integrativa de Literatura (RIL), la cual fue empleada para la presente investigación en razón de su uso innovador para las ciencias sociales. Después se analiza la evolución histórica de la manipulación fotográfica hasta llegar a los orígenes de los *deepfakes*. Posteriormente, se esclarece la naturaleza de los *deepfakes*, revelando el vasto número de definiciones que se han tejido en torno a este tipo de tecnología, así como también sus características. Luego se examinan las implicancias de los *deepfakes* en diversos sectores, haciendo un énfasis especial en el ámbito jurídico. Finalmente, se exponen las problemáticas suscitadas de la injerencia de los *deepfakes* a bordo del derecho probatorio, tanto en *common law* como en *civil law*.

## METODOLOGÍA

La ejecución del presente estudio se realizó a través de la metodología denominada Revisión Integrativa de Literatura (RIL), debido a que esta no solo permite la adecuada integración de diversos datos, tales como cuantitativos, cualitativos u otras formas de información eficaz, sino que también hace viable la superación del proceso de

análisis y síntesis de los descubrimientos de la investigación primaria. Consecuentemente se puede alcanzar una comprensión más amplia del fenómeno analizado (Lubbe, Ham-Baloyi y Smit, 2020, p. 308).

Asimismo, si bien existen diversas propuestas respecto a las fases que deben componer la RIL, esta investigación se alinea a las propuestas de Russell (2005) y Torracó (2016), las cuales gozan de gran similitud, que son: 1) formulación del problema; 2) recolección de data; 3) evaluación de la data; 4) análisis de la data; y 5) interpretación y presentación de los resultados.

En ese sentido, durante la primera fase se estableció que el objetivo de la pesquisa sería analizar el probable impacto de la tecnología *deepfake* en el derecho probatorio; posteriormente, durante la segunda fase, se recolectaron artículos científicos provenientes principalmente de *Scopus* y *Web of Science*. Cabe añadir que, debido a la temática poco estudiada, se tuvo que considerar estudios interdisciplinarios que tenían como objeto de estudio a los ultrafalsos. Es así que, durante la tercera y cuarta fases, se filtraron todos los documentos recolectados, seleccionando únicamente aquellos que respondieran a las siguientes cuestiones: ¿qué son los *deepfakes*?, ¿cuáles son las implicancias provenientes de los *deepfakes* en la sociedad? y ¿qué impacto tienen estas tecnologías en el derecho probatorio? Finalmente, durante la quinta

fase, se analizó toda la información seleccionada y se evidenció que los *deepfakes* no solo presentarían dificultades para el ámbito probatorio en *civil law*, sino también en *common law*.

### **BREVE HISTORIA DE LOS “DEEP-FAKES”: DE LOS ESPÍRITUS MUMLER A REDDIT**

A pesar de que, desde su origen en 1839, a la fotografía se le consideró capaz de expresar la verdad más absoluta un cúmulo de sujetos reconocieron al poco tiempo de la introducción de este artefacto, la potencialidad de explotar las ideas inestables de la gente en torno a la naturaleza de la realidad y la verdad (Tucher, 2022, p. 92). Siendo así, el autor reconoce como uno de los primeros manipuladores de imágenes a William H. Mumler, un exgrabador que entre los años de 1860 y 1870 se hizo conocido por ser un médium fotógrafo de espíritus, ya que aseguraba que podía tomar fotografías de personas muertas<sup>2</sup> (Figura 1). Cabe resaltar que el actuar de Mumler generó diversas burlas y denuncias, pues se estaba atentando contra la integridad del oficio de los fotógrafos profesionales y la razón y la lógica; por consiguiente, en 1869, Mumler, acusado por fraude y hur-

2 El *modus operandi* de Mumler se basaba en colocar al modelo vivo en una pose habitual, para que posteriormente mediante la aplicación de técnicas de manipulación dentro del cuarto oscuro produjera una imagen dentro de la cual se hallara claramente la figura fantasmal de un ser querido detrás o sobre el hombro del sujeto vivo (Tucher, 2022, p. 92).

to, fue llevado a los tribunales de Manhattan para una audiencia preliminar; sin embargo, el juez no pudo enviar el caso a la siguiente instancia en razón de que, si bien poseía cierto convencimiento de que el acusado había cometido engaños, la fiscalía, a pesar de haber investigado el caso por tres semanas, no pudo demostrar cómo Mumler había generado las fotografías.

Por su parte, Jitendra Sharma y Rohita Sharma (2017), analizando el impacto de la manipulación fotográfica en la historia

del fotoperiodismo, encuentran como casos más resaltantes la fotografía de Abraham Lincoln, la cual fue compuesta por el cuerpo del político John Calhoun y la cabeza de Lincoln (Figura 2); la foto de Joseph Stalin, en la cual se suprimió un comunista popular con fines propagandísticos, debido al colapso de la Unión Soviética (Figura 3); y la fotografía de Benito Mussolini, que fue modificada eliminando al cuidador de caballos en la foto original, para que de este modo se proyecte una imagen más heroica (Figura 4), entre otras.

**Figura 1** Ejemplos de fotografías tomadas por William H. Mumler



*Nota:* colección de fotografías del Centro Getty.

**Fuente:** Waldorf (2021).

**Figura 2** Manipulación de la fotografía de Abraham Lincoln

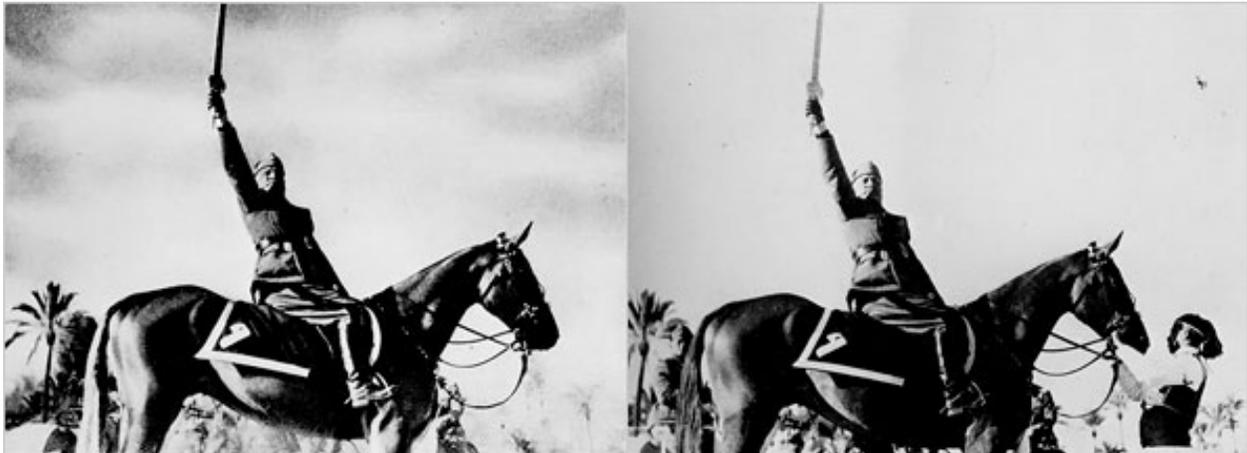


**Fuente:** Waters (2017).

**Figura 3** Foto modificada de Joseph Stalin con fines políticos



**Fuente:** Rare Historical Photos (2021).

**Figura 4** Fotografía de Benito Mussolini alterada

**Fuente:** *The New York Times* (2009).

Si bien se puede advertir que ya era posible poder adulterar fotografías mediante el uso de procedimientos analógicos o químicos de laboratorio, con el advenimiento y mejora de tecnologías, como programas de edición fotográfica, se produjo una evolución exponencial en la manipulación fotográfica, que estuvo enfocada principalmente en el ámbito de la publicidad (Bravo, 2017, p. 207). Es de menester indicar que este salto cualitativo generó casos de fraude polémicos, tales como la instantánea de las pirámides de Giza, en la cual Galen Gordon tuvo que reducir la distancia entre las pirámides de manera digital, para que pudiera encajar en el diseño de la revista *National Geographic*; o el caso del fotógrafo Souvid Datta, quien admitió que había manipulado diversas fotos y que, además, había empleado los trabajos fotográficos de sus colegas como insumo para producir otras instantáneas (Sharma y Sharma, 2017, p. 92).

Posteriormente, con la revolución generada por la IA dentro del plano artístico, especialmente por el área denominada visión por computadora, se dio origen a lo que conocemos como *deepfakes*; de hecho, se pueden identificar cuatro eventos clave que suscitaron el arranque de este tipo de tecnología: el primero fue el programa Rewrite,<sup>3</sup> desarrollado por Christoph Bregler, Michele Covell y Malcolm Slaney en 1997, cuya caracterización se puede encontrar en el artículo denominado “Video Rewrite: Driving Visual Speech with Audio”; asimismo, el segundo fue el desarrollo de los modelos

<sup>3</sup> El programa Rewrite utiliza metraje real para generar el video de una persona pronunciando palabras no realizadas en el metraje original; es por ello que se sugiere su empleo para el doblaje de películas (Bregler et ál., 1997, p. 353). Además, los autores acotan que Video Rewrite etiqueta los fonemas en la nueva pista de audio y los datos de entrenamiento; después reordena las imágenes de la boca del video de entrenamiento, para que de este modo exista coincidencia con la secuencia de fonemas de la nueva pista de audio.

de apariencia activa<sup>4</sup> producido por Timothy Cootes, Gareth Edwards y Christopher Taylor en 2001, y descrito en la investigación nombrada “Active Appearance Models”; seguidamente, el tercero fue el proyecto Face2Face<sup>5</sup> creado por la Universidad Técnica de Múnich; y, finalmente, en cuarto lugar, el proyecto Synthesizing Obama<sup>6</sup> de la Universidad de Washington (Song, 2019).

Fuera del ámbito académico fue a finales de 2017 cuando la tecnología *deepfake* tomó mayor popularidad en razón de que un usuario de Reddit —llamado “Deep-

fakes”— publicó un gran número de videos, en los que se podía apreciar que en diversas escenas pornográficas se había intercambiado el rostro de diversas actrices de Hollywood por intérpretes del cine para adultos; además, unos meses después se compartió en Reddit una aplicación que permitía que cualquier usuario pudiera crear su video mediante el intercambio de rostros (De Ruiter, 2021, p. 1314). Asimismo, el hecho de que la producción y uso de *deepfakes* hayan dejado el plano de la complejidad hizo que los cibernautas no tardaran mucho en abusar de esta técnica para crear imágenes hiperrealistas, que si bien en una primera instancia se enfocaban únicamente al escenario pornográfico, la atención migraría a otros sectores como el político y el científico, entre otros; llegando a un punto donde no solo se menoscabara el honor de los afectados, sino que se llegó a poner en peligro bienes vitales para la humanidad, como la verdad.

En suma, a pesar de que la manipulación de imágenes es inherente al nacimiento de la cámara fotográfica y es muy probable que vivamos entre recuerdos sintéticos, lo cierto es que la tecnología *deepfake*, debido a su disposición y fácil uso, sumado a la capacidad de las redes sociales para la publicación de información a escala de una

4 A través del artículo se introduce un método que tiene como objetivo principal generar coincidencias entre imágenes y modelos estadísticos de apariencia; más aún, los autores desarrollan la construcción del Algoritmo de Coincidencia Iterativo Eficiente considerando la relación existente entre los errores de imagen inducidos y las perturbaciones en los parámetros del modelo (Cootes et ál., 2001, p. 681).

5 El presente enfoque se configura como uno de los primeros sistemas de recreación facial en tiempo real que solo demandaría una entrada RGB monocolor; también la configuración en tiempo real posibilitaría la animación de secuencias heredadas en vivo (Thies et ál., 2019, p. 1). Aunado a ello, los autores revelan que el objetivo de Face2Face es animar las expresiones faciales del video de destino por un actor fuente y de este modo volver a exportar el video de salida manipulado de manera fotorrealista.

6 La tarea fundamental era generar videos fotorrealistas de Barack Obama, partiendo de archivos de videos y voz. Cabe mencionar que el expresidente Obama fue elegido como sujeto de prueba, debido a que existe un bagaje considerable de videos colgados en red de discursos del exmandatario (alrededor de diecisiete horas y casi dos millones de fotografías), los cuales pertenecen al dominio público (Suwajanakorn et ál., 2017, p. 1).

manera nunca antes vista, hace que a la presente técnica se le asocien potenciales usos y problemáticas, por lo que se hace imperioso el análisis de los *deepfakes*.

### **ANATOMÍA DE LA TECNOLOGÍA “DEEPFAKE”: DISECCIONANDO LA REALIDAD ARTIFICIAL**

A pesar de que la palabra *deepfake* es producto de la combinación de *deep* (haciendo referencia a *deep learning*) y *fake*, como consecuencia de la edad temprana en la que aún se encuentra este tipo de tecnología, se ha podido evidenciar una suerte de evolución respecto a su conceptualización por parte de diversos sectores; tal es así que encontramos definiciones básicas como la del diccionario de Oxford (2024), mediante la cual se estipula que los *deepfakes* son videos de una persona cuya apariencia fue digitalmente alterada para que pueda parecer otra persona, empero se cree que este enfoque no posee una adecuada delimitación, debido a que se restringe únicamente a videos, dejando de lado otras técnicas como la fotografía y el audio.

Asimismo, encontramos otro grupo de conceptualizaciones más detalladas como la de Somers (2020), quien, basándose en su finalidad y el género-especie, afirma que un *deepfake* puede ser entendido como un tipo de medio sintético<sup>7</sup> en el que se

intercambia a una persona en una imagen o video por otra. También, Altuncu, Franqueira y Li (2022, p. 23) establecen que un *deepfake* es un tipo de medio sintético en el que se intercambia a una persona en una imagen o video por otra y es usualmente empleado para hacer referencia al desarrollo de nuevos medios de carácter sintético o la manipulación de medios, tales como audio, video o imagen mediante los enfoques basados en *deep learning* (Altuncu et ál., 2022, p. 23). Por su parte, autores como Gil et ál. (2023, p. 11295), enfocándose en el procedimiento de generación de *deepfakes*, advierten que los ultrafalsos<sup>8</sup> son el resultado de aplicaciones de IA, que combinan, fusionan, reemplazan y superponen videoclips o imágenes para generar videos falsos que simulen ser reales.

Por otro lado, autores como Jacobsen (2023, p. 1106), buscando analizar la naturaleza de este tipo disruptivo de técnica, conceptualiza los *deepfakes* como “un conjunto de tensiones diferenciales en la sociedad”; en donde por conjunto se debe entender que los *deepfakes* nunca serán una sola cosa, debido a que son una suerte de fusión compleja entre datos, medios, algoritmos, contextos socioculturales y políticos, incentivos financieros y daños sociales reales y percibidos, entre otros.

<sup>7</sup> IA, que se caracterizan principalmente por su grado de inmersión y realismo, lo que hace a la mayoría de usuarios se les imposibilite el poder diferenciarlos de otros medios del mundo real (Van Rijmenam, 2022).

<sup>8</sup> Otra forma de llamar a los *deepfakes* en Iberoamérica.

<sup>7</sup> El término medios sintéticos abarca una innovadora forma de medios virtuales generados mediante

En esa misma línea, el autor explicando la relación existente entre los *deepfakes* y las tensiones diferenciales indica que este tipo de tecnología representa múltiples fuerzas en competencia dentro de la sociedad, es decir, rupturas, perpetuaciones, continuidades y discontinuidades; en ese sentido, los *deepfakes* representarían tanto la ruptura como la continuación de la variada política de la imagen en el mundo social.

Por ello, considerando el universo de conceptos sobre los ultrafalsos, se cree que a los *deepfakes* se les puede definir desde dos aristas: en primer lugar, como tecnología al ser un procedimiento que emplea *deep learning* para la generación de imágenes artificiales. En segundo lugar, como producto, ya que es un tipo de medio sintético creado mediante IA en el que se busca modificar o manipular documentos audiovisuales reales o ficticios.

Asimismo, según Naitali et ál. (2023) se han llegado a identificar cinco tipos de manipulación de *deepfakes*: 1) *síntesis facial*, que implica la generación de un rostro que no existe en el mundo real; 2) *manipulación de atributos*, que busca generar modificaciones en una determinada área descollante para cambiar la apariencia facial a través del retoque en la piel, quitándose o poniéndose anteojos, o incluso el cambio de género o edad; 3) *intercambio de caras*, la cual goza de mayor popularidad y se entiende como la sustitución del rostro

de una imagen o video, el cual es llamado fuente, por el rostro de otra persona denominado objetivo; 4) *recreación facial*: es aquella técnica enfocada en la alteración de las expresiones faciales de una persona en un video y, del mismo modo, se buscaría intercambiar la expresión original de un determinado sujeto por la de otra persona; finalmente, 5) *sincronización de labios*, mediante la cual se pretende desarrollar una cara objetivo que parezca real y que, además, mantenga una sincronización con entradas de audio o texto determinadas (Naitali et ál., 2023, p. 4).

Respecto a la funcionalidad, se evidencia que los *deepfakes* se configuran como el resultado de redes generativas adversarias (GAN); en otras palabras, dos redes neuronales artificiales cooperarían para que se puedan desarrollar medios de apariencia real. Cabe resaltar que estas redes recibirán el nombre de discriminador y generador, y que el procedimiento para la generación de un *deepfake* consiste en que la red generadora busca desarrollar un cúmulo de muestras que posean un alto grado de calidad, para que de este modo pueda engañar a la red discriminadora, la cual se encarga de analizar si las muestras que vea son reales; siendo así, se da una mejora constante buscando de esta manera el perfeccionamiento (Westerlund, 2019, p. 41). El autor también revela que si bien, actualmente las GAN necesitan apreciar miles de imágenes de un sujeto para

generar un nuevo retrato sin ser una copia exacta del original, el desarrollo exponencial hace que esta tecnología en un futuro pueda generar un video teniendo como material de entrenamiento únicamente una fotografía.

En definitiva, el estado pueril en el que se encuentra este tipo de tecnología no solo se evidencia en el universo de conceptos que se hacen a su alrededor, sino, además, en el hecho de que hasta ahora no se han delimitado sus potenciales usos. Es así que se hace necesario analizar el conjunto de usos y problemas que podrían desprenderse de esta innovadora tecnología.

### **POTENCIALES USOS Y ABUSOS DE LOS “DEEPFAKES”: ANALIZANDO EL IMPACTO DE LOS ESPEJISMOS 4.0**

A pesar de que en un inicio los *deepfakes* solo se enfocaban en el ámbito pornográfico, durante los últimos años la atención de miles de fanáticos a esta tecnología se desplazó a otros campos como el artístico o el político; y es que los cibernautas no demoraron en producir imágenes en las que se podía ver al expresidente Donald Trump pelear contra un policía o al papa Francisco con un chaquetón blanco de la marca Balenciaga creado a través de Midjourney. Paralelamente, un bagaje ilimitado de videos falsos se gestionaba en el espacio digital, de los cuales resaltan el de Jerry Seinfeld actuando en la película *Pulp fiction*, de Quentin Tarantino, o el vi-

deo sintético producto de la colaboración entre Malaria No More, RG/A y Ridley Scott Associates, a través del cual se puede ver a David Beckham hablando nueve idiomas distintos (Sodji, 2023).

De lo mencionado anteriormente, observamos que si bien en un comienzo se creaban imágenes manipuladas mediante IA para hacer comedia o atentar contra el honor en los casos de la pornografía artificial, el avance exponencial de la IA, junto con la malicia de diversos sujetos, ha generado un estado de indefensión para un cúmulo de bienes jurídicos pertenecientes a diversos sectores; es por esta razón que se analizará el impacto de los *deepfakes* en diversos sectores:

- **Ámbito empresarial:** acorde a Mustak, al analizar el engaño al mercado, a través de los *deepfakes*, se identifican las *potenciales amenazas al sector empresarial*, dentro de las cuales podemos distinguir la posibilidad de que actividades despectivas, como el sabotaje, la difamación o la creación de reseñas falsas sobre productos promovidas por la proliferación de ultrafalsos, afecten directamente la marca e imagen empresarial de manera íntegra, lo cual llevará eventualmente a la pérdida de confianza por parte de los *stakeholders* (2023, p. 6); y las *potenciales amenazas a los consumidores*, en las que se halla la probabilidad de que aquellos clientes que carezcan de competencias cognitivas para

la detección de imágenes manipuladas no solo se dejen convencer por información falsa, sino que también tengan predisposición a promulgar la información digital inexacta. De hecho, los autores hallaron que existen sectores demográficos propensos a ser víctimas de engaño por contenido falso, debido a que en su estudio encontraron que los usuarios de Facebook mayores de 65 años llegaron a compartir casi 7 veces más artículos que contenían *fake news* que un grupo de edad más joven. Además, los *deepfakes* generarían diversas complicaciones durante el procedimiento de búsqueda de empleo, debido a que el 90 % de empleadores, aproximadamente, se basan en los resultados de búsquedas en red para la adquisición de talento humano y, como consecuencia, un gran número de solicitantes se hallan perjudicados por imágenes inapropiadas encontradas durante la búsqueda; por ende, la creación de imágenes falsas disminuyen considerablemente las perspectivas laborales de las personas. Cabe indicar que la información falsa digital, se encuentra relacionada con el efecto de la tercera persona, es decir, la tendencia de los sujetos a reparar que los demás están mucho más influenciados por los mensajes provenientes de los medios de comunicación que ellos mismos (Tsay-Vogel, 2020, p. 1).

Asimismo, respecto a las ventajas que se desprenden del uso de los *deepfakes* a bordo del sector empresarial, se encuentran

el diseño y la realización de campañas de *marketing* a bajos costos; la posibilidad de crear embajadores digitales de marca, como Lil Miquela; la capacidad de desarrollar nuevas ofertas y modelos de negocio, ya que se podría generar una innovadora forma de personalizar productos, servicios o marcas, tal y como lo hizo el canal de televisión surcoreano MBN, el cual presentó al público un *deepfake* del conductor Kim Joo-Ha; y, finalmente, se lograría generar valor en los sistemas ciberfísicos, pues se podrían generar humanos digitales, los cuales gozarían de un alto grado de realidad y que, además, alcanzarían a interactuar con el entorno, es decir, se constituiría la *deep resurrection* o resurrección profunda, tal y como se puede apreciar en el sector turístico en el Museo de Salvador Dalí en San Petersburgo, en donde los espectadores pueden entablar una interacción con el artista (Mustak et ál., 2023, p. 10).

- **Ámbito científico:** si bien las actividades académico-científicas se hallan enmarcadas en un conjunto de reglas para fomentar la conducta ética en las pesquisas en general, el auge de los *deepfakes* promueve una mala conducta en la investigación de alta calidad y en la fabricación de datos para la manipulación de resultados, específicamente la de fraude de imágenes, la cual se ejecuta cuando se manipulan imágenes con el objetivo de distorsionar

sus significados. Conviene aclarar que el acto constituiría una falta grave en razón de que las ilustraciones desempeñan un papel fundamental al momento de transmitir los resultados científicos; asimismo, este tipo de fraude no solo deterioraría la confianza dentro de las comunidades de investigación, sino que también afectaría la reputación que posee la ciencia frente a la sociedad (Wang et ál., 2022, p. 1). En esa misma línea, Wolofsky (2014, p. 96) establece que la fabricación de ilustraciones digitales lleva mucho tiempo en el ámbito académico<sup>9</sup> y que como respuesta a la presente problemática aparecieron innumerables guías éticas; a pesar de ello, se cree que las guías éticas no bastarán para aplacar esta actividad, es más, el presente estudio considera que la única forma de proteger la verdad científica sería mediante la creación de algoritmos de detección de *deepfakes*, los cuales evolucionan conjuntamente con estas malas praxis digitales. Asimismo, se hace imperioso advertir que existe una gran probabilidad de que los investigadores tiendan a basarse en *fake news* para la elaboración de artículos científicos, lo que conllevaría eventualmente a que el producto final de la pesqui-

sa mantenga un sesgo; y, además, lo más alarmante es que todos los investigadores que usen ese trabajo como fuente generen desinformación masiva.

En contraparte, los *deepfakes* podrían servir de apoyo para desarrollar investigaciones experimentales dentro del campo de las ciencias sociales, ya que a diferencia de las investigaciones que utilizan robots o avatares, los *deepfakes* no causarían problemas como las del valle inquietante, debido a la sensación agradable que pueden ofrecer al público (Eberl, Kühn y Wolbring, 2022, p. 3). Ahora bien, los autores demandan una diferenciación en torno al empleo de los *deepfakes*, ya que si bien el uso de esta tecnología para engañar a las personas es un acto moralmente reprochable, el servirse de ultrafalsos como instrumentos científicos para generar condiciones experimentales óptimas, se establece como una oportunidad para un buen uso de estas tecnologías disruptivas.

- **Ámbito social:** los *deepfakes* tienen la particularidad de generar contenidos falsos bajo una apariencia real; siendo así, este tipo de técnicas sofisticadas se configuran como un arma poderosa para la construcción de *fake news* o noticias falsas (Molina y Berenguel, 2022, p. 1). Más aún, instituciones como el Colegio Técnico del Área de Milwaukee MAT han llegado a concebir a las noticias falsas como *deepfakes* con esteroides (2024). Y si bien, podemos

9 Nicholas Wade (2006) indica que en *The Journal of Cell Biology* entre 2002 y 2006, como producto de una prueba, se halló que el 25 % de todos los artículos aceptados tienen, al menos, una ilustración manipulada, lo cual atentaría contra las políticas de la institución; y si bien, la mayoría de los casos se resolvieron debido a que los científicos enviaron las imágenes reales, existe un 1 % de casos en los que se encontró fraude.

reconocer que, al igual que las manipulaciones fotográficas, el advenimiento de las noticias falsas va ligado a la creación de la imprenta, el avance desmedido de la IA, junto con el vasto alcance de las redes sociales en general, hace que los *deepfakes* sean ubicuos. Asimismo, aun cuando la empresa de analizar el impacto de las *fake news* resulte engorrosa, lo cierto es que se ha generado una industria de medios de noticias falsas; siendo así, existe un grupo de personas que la financian, porque evidentemente se obtienen beneficios de la operatividad de la cadena de valor de las noticias falsas<sup>10</sup> (Rodny-Gumede, 2018, p. 206). Por otro lado, se logró identificar que las noticias falsas girarían en torno a dos agendas: comerciales y políticas; a tal efecto, se hace necesario analizar el campo sobre el que se desarrolla una noticia falsa, ya que tanto el objetivo como las consecuencias son distintos, dado que difiere mucho el sitio que solo propaga chismes falsos para generar un mayor tráfico del sitio que busca perjudicar y promover

un determinado partido político; si bien ambas agendas poseen determinadas características no son excluyentes entre sí (Rodny-Gumede, 2018).

En ese sentido, si bien hasta ahora se han evidenciado los potenciales efectos sociales nocivos que podría desplegar la proliferación de los *deepfakes* en general es probable que nos encontremos viviendo el macroefecto de los abusos de los ultrafalsos, es decir, nos hallamos en el preludio de lo que es conocido como la era de la posverdad, la cual es definida por el Diccionario de Oxford (2024) como “Relacionado con circunstancias en las que las personas responden más a sentimientos y creencias que a hechos”; así como también por autores como McIntyre (2018, p. 12) como una forma de supremacía ideológica caracterizada por que sus practicantes tienden a coaccionar a alguien a que crea en un algo, a pesar de que no exista fundamento alguno; concibiendo como principal afectada a la verdad.

· **Ámbito jurídico:** se considera que probablemente el sector legal se vea más afectado no solo por el hecho de que la generación de ultrafalsos *per se* conseguiría potencialmente configurarse como un acto delictivo, el cual podría tener diversas variaciones dependiendo del bien jurídico perjudicado, como el honor, la democracia o la verdad, sino que también resulta necesario analizar la forma en que inteli-

10 BuzzFeed, empresa enfocada a los medios de comunicación en el internet estadounidense, investigando el origen de la noticia sobre el apoyo del papa Francisco a Donald Trump en las elecciones presidenciales de Estados Unidos, descubrieron que el sitio web donde se originó la noticia falsa formaría parte de una red macro conformada por 43 sitios web, aproximadamente, los cuales publicaron 750 *fake news* en 2016; cabe resaltar que ninguno de los sitios se hallaría asociado a una empresa matriz o propietaria (Rodny-Gumede, 2018, p. 207). Además, se pudo verificar que diversas noticias falsas tenían por objetivo el favorecer a determinados candidatos en las elecciones estadounidenses.

gencias artificiales generativas provocan diversas cuestiones respecto a derechos de autor en la medida en que los sistemas inteligentes emplean durante su entrenamiento imágenes con licencia para la producción de archivos. Sin embargo, se cree que el mayor peligro se centra en la injerencia de los *deepfakes* dentro del derecho probatorio, ya que si bien la prueba electrónica o e-prueba goza de un marco regulatorio procesal innovador, la sola presencia de ultrafalsos a bordo de un litigio plantea diversos dilemas procesales, entre los que se encuentran la duda degenerativa en las partes procesales hacia cualquier medio probatorio presentado o el estado de indefensión en el que se encuentra la verdad procesal, entre otros.

Es así que a continuación se analizará el tratamiento jurídico de los *deepfakes* por parte de diversos Estados, así como el impacto que tienen los ultrafalsos respecto al derecho probatorio, tanto en *common law* como en *civil law*, para que de este modo se pueda esbozar una probable respuesta a esta inadvertida problemática digital.

### **¡OBJECCIÓN, TODO ES FALSO!: EL DERECHO Y LOS MEDIOS SINTÉTICOS**

El encuentro entre los *deepfakes* y el derecho ha generado dos áreas críticas que demandan el estudio de académicos y técnicos jurídicos: el tratamiento legal de la creación de *deepfakes* como delitos; y la in-

trusión de los *deepfakes* en el derecho probatorio.

A pesar de que a escala internacional la regulación de los ultrafalsos *extra processum* es muy reciente, resulta imperioso el analizar la forma en que abordan el problema de los *deepfakes* como actos delictivos; es así que a continuación se revisarán los instrumentos legales diseñados por diversos países para enfrentar esta problemática digital.

Continente europeo: en el Reino Unido, durante 2024, se presentó un proyecto de ley con el cual se pretende sancionar a aquellos sujetos que crean pornografía *deepfake* con una multa ilimitada y la posibilidad de una pena privativa de la libertad en caso de que las imágenes circulen por la red de manera masiva (Gov.UK, 2024). Del mismo modo, la Comisión de la Unión Europea dentro de sus diversos proyectos procura ampliar la definición de abuso infantil para criminalizar la creación y divulgación de ultrafalsos, así como la retransmisión en directo de actos abusivos (Thissen, 2024).

Continente asiático: en China, el 10 de enero de 2023, entraron en vigencia las Disposiciones Administrativas de China sobre Síntesis Profunda en los Servicios de Información Basados en Internet (Disposiciones sobre Síntesis Profunda), a tra-

vés de las cuales se establecen responsabilidades para los proveedores de servicios de síntesis profunda (SSP) y los soportes técnicos de SSP sobre la protección de información y datos personales, la seguridad técnica, la gestión y etiquetado de contenidos, y la transparencia, entre otros (Zhang, 2023). Asimismo, en Corea del Sur tras la revisión de la Ley de Elección de Funcionarios Públicos, se prohibieron los *deepfakes* relacionados con los candidatos políticos por un periodo de 90 días antes de las elecciones; cabe señalar que la sanción podía ascender a 7 años de pena privativa de la libertad o una multa de hasta \$ 37 500 dólares (Lee, 2024).

Continente americano: en la provincia de Columbia Británica (Canadá), el 29 de enero de 2024 entró en vigencia la Ley de Protección de Imágenes Íntimas, norma que abarcaría los ultrafalsos por medio de la que cualquier persona cuyas imágenes íntimas se hayan distribuido sin consentimiento, o incluso, si es que hubo consentimiento, pero después se haya revocado, ahora cuenta con diversos recursos legales, ya que el Tribunal de Resolución Civil se encuentra facultado para ordenar la efectiva eliminación de imágenes y, además, sancionar a los responsables con multas de hasta \$ 500 dólares por día para las personas naturales y \$ 5000 dólares por día para aquellos sitios web que no cumplan con un mandato de elimi-

nación del material (Keith, 2024). Por su parte, Estados Unidos de América en sus primeros intentos por regular este tipo de tecnología promovió a finales de 2019 la Ley de Autorización de Defensa Nacional de 2020, a través de la cual se exige que la Comunidad de Inteligencia Norteamericana analice el uso de *deepfakes* como arma de terceros internacionales, como Rusia o China, y que se emita un informe cada vez que un país del extranjero desarrolle un ultrafalso para afectar las elecciones estadounidenses; asimismo, se puede advertir esfuerzos de sus estados de manera independiente, como el estado de Virginia, en donde se enmendó durante 2019 su Ley de Pornografía por Venganza, para que de este modo se penalice a todo aquel sujeto que por medio de ultrafalsos con carácter sexual coaccionen, acosen o intimiden a otra persona; en esa misma línea, el estado de Texas durante el mismo año modificó su Código Electoral para penalizar la creación y distribución de *deepfakes* enfocados a afectar a un candidato político; finalmente, el estado de California promulgó un nuevo derecho de acción civil para que las víctimas de pornografía ultrafalsa puedan buscar una reparación por daños (Delfino, 2023).

En suma, se puede evidenciar que un grupo de países han aunado esfuerzos para regular los ultrafalsos, sobre todo en el ámbito político, debido al estado de vulneración en el que se encuentra la de-

mocracia promovido por el avance de la IA. Aunado a ello, se aprecia que únicamente algunas naciones desarrolladas han procurado o están en vías de regular esta actividad cibernética, de lo que se desprende la generación de una brecha producto de los beneficios que solo podrán gozar aquellos países que regulan estos ciberdelitos; además, se manifiesta nuevamente el divorcio existente entre el derecho y la realidad, debido al crecimiento exponencial de las tecnologías frente al avance aritmético de los sistemas normativos.

Por otro lado, al analizar las implicancias derivadas de la presentación de *deepfakes* como medios probatorios *in processus*, se considerará tanto el sistema *common law* como el *civil law* y la problemática de los ultrafalsos y el peritaje.

#### *Los deepfakes a bordo del sistema anglosajón*

Considerando el estudio de Delfino (2023, pp. 308-312), se extrae que los desafíos que plantean los *deepfakes* en los tribunales son: demostrar si una imagen o audio presentados como prueba son legítimos; enfrentar la desconfianza de las partes procesales hacia la autenticidad de todas las imágenes digitales o pruebas de audio; y responder a la defensa de los *deepfakes* como también al alegato de que pruebas genuinas son ultrafalsas. Como respuesta a estos cuestionamientos el autor identifica diversos grupos dentro de los cuales

se encuentra un sector optimista, el cual sostiene que tanto las Reglas Federales de la Prueba como las normas prácticas son lo suficientemente eficientes como para abordar esta problemática; y un sector pesimista, que demanda diversas modificaciones a través de enmiendas respecto a la Regla 901, mediante la cual se regula la autenticación de las pruebas digitales, y, además, afirma que existe la necesidad de exigir más pruebas circunstanciales para corroborar evidencias como las de video; y, finalmente, un grupo intermedio que manifiesta la necesidad de que los tribunales redefinan la cantidad y calidad de las pruebas circunstanciales, para que de esta forma el jurado razonablemente estime la autenticidad de las pruebas digitales de carácter audiovisual acorde a cada teoría de derecho consuetudinario, es decir, ya sea la Teoría de Comunicación Pictórica<sup>11</sup> o la del Testigo Silencioso.<sup>12</sup>

11 Por Teoría de Comunicación Pictórica o Testimonio Pictórico, se entiende que la admisibilidad de la evidencia audiovisual depende de la posibilidad de que un testigo exponga ante un jurado que la evidencia es una representación precisa y justa de lo ocurrido (Delfino, 2023, p. 327). Cabe indicar que la presente teoría se hallaría sustentada por la Regla 901 "Autenticación o identificación de pruebas" dentro del primer ejemplo propuesto (b)(1), "testimonio de un testigo con conocimiento: Testimonio de que un ítem es lo que se dice que es". Delfino sostiene que la Teoría del Testimonio Pictórico se basa en que tanto el video como la fotografía son una representación gráfica del testimonio oral; en ese sentido, se requiere la validación por parte de un testigo, el cual debe poseer conocimiento personal del tema para hacer confiables los hechos presentados en las pruebas.

12 La Teoría del Testigo Silencioso, por su parte, se basa en la Regla 901, específicamente en el punto (b)(9), mediante el cual se admite cualquier prueba

Es así que se concuerda en parte con Delfino (2023), quien indica que ninguna regla sobre la prueba en el sistema anglosajón es adecuada para dar respuesta a las cuestiones generadas alrededor de los *deepfakes* en un tribunal; consecuentemente en primer lugar se debe dejar de emplear la Teoría del Testigo Silencioso, ya que si bien la mayoría de órganos jurisdiccionales norteamericanos se inclinan por aplicar la presente teoría en razón de la fe depositada por parte de la sociedad hacia pruebas como los videos de cámaras de seguridad, confiar en estas pruebas en la era de los *deepfakes* solo llevaría a un error generalizado hacia los tribunales en general; en segundo lugar, el autor suscita que se deben replantear los papeles entre el tribunal y el jurado, debido a que en el presente el tribunal de primera instancia se encarga de resolver cuestiones de autenticidad de las pruebas y considerando que los miembros del jurado carecen de conocimientos técnico-digitales, se hace necesario que el juez tenga el monopolio de la autenticación de pruebas, pues este magistrado gozaría de mayores conocimientos técnicos para afrontar los ultrafalsos; y en tercer lugar, el autor bosqueja una modificación normativa a la Regla 901 mediante la que se indica que sin afec-

tar lo estipulado en la subdivisión (a) “el proponente debe presentar evidencia de que el elemento es lo que el proponente afirma”; por consiguiente, el producto de estas tres propuestas harían viable una autenticación eficiente de la prueba y, además, el tribunal enfrentaría eficientemente la posibilidad de que los abogados se aprovechen de las dudas tejidas sobre la autenticidad de las pruebas, solicitándoles que no presenten tales argumentos.

Si bien Delfino (2023) conjetura que los jueces poseen los conocimientos suficientes para autenticar pruebas digitales, esto sería un error, ya que si bien los órganos jurisdiccionales poseen un cúmulo de conocimientos fruto de su profesionalismo, el avance desproporcional de la tecnología superaría considerablemente su entendimiento sobre los ultrafalsos; por tanto, opinamos que el juez deberá tener una relación de dependencia con los peritos informáticos para encarar el dilema de los *deepfakes*.

#### *Los deepfakes y el civil law*

En primer lugar, debido al desarrollo de los medios de comunicación un gran número de personas pueden tener conocimiento de un suceso en tiempo real, por lo que estos hechos gozarían de la calidad de notorios (Camacho, 1998). Es de menester reconocer que la doctrina exime de prueba al hecho notorio, a pesar de que las partes

---

audiovisual sin la necesidad de un testigo calificado; es así que las partes procesales confían en que el proceso mediante el cual se produce la prueba es idóneo (Delfino, 2023, p. 330).

no lo acepten, pues lo primordial es que la notoriedad sea clara para el juez (Devis, 1984, p. 109). Sin embargo, al encontrar a damnificados por los malos usos de los medios sintéticos, se hace inviable la posibilidad de constituir un hecho notorio.

En segundo lugar, respecto al peritaje de las pruebas electrónicas, Cano (2024) indica que al vaticinar que los jueces carecen de conocimientos para diferenciar un *deepfake* de un documento audiovisual auténtico, se hace imperiosa la necesidad de un perito, el cual deberá emplear tecnologías basadas en IA para poder llevar a cabo la detección de ultrafalsos, lo cual generará cuestiones, debido a que no existen estándares de calidad para la realización de esa pericia y que, del mismo modo, la pericia aumenta los costos del proceso, por lo que es probable que en muchos casos de escasa cuantía, no se puedan llevar a cabo estos peritajes; además, señala que existe la posibilidad de que el apoyo de la IA para identificar los *deepfakes* sea insuficiente, pues ya se ha visto que muchos sistemas no logran identificar los ultrafalsos. También, Cano (2024) sostiene que el hecho de que la IA sea la que haga el dictamen involucra que el perito rinda cuentas en torno al resultado que realizó una herramienta, mas no su persona, lo que lleva a una paradoja desde el punto de vista de contradicción; ahora bien, al analizar la calidad científica de las pericias de *deepfakes* a la luz de los

criterios Daubert, se puede apreciar que el procedimiento de la IA para detección de ultrafalsos al gozar de opacidad y, además, errores, no cumpliría con los dos primeros criterios: 1) que la técnica utilizada por el perito ha sido probada suficientemente frente a errores; y 2) que la técnica ha sido revisada por otros científicos y, en su caso, ha sido publicada (Nieva, 2018, p. 15).

En suma, dentro de la constelación de problemáticas que presentan los *deepfakes* en el ámbito probatorio encontramos respecto al *common law*, la insuficiencia del Principio del Testigo Silencioso para la admisión de pruebas digitales por parte de los tribunales norteamericanos en razón del advenimiento de los ultrafalsos, y la potencial dificultad que tiene el jurado frente a la valoración de la pruebas electrónicas, debido a la carencia de conocimientos técnico-digitales; asimismo, en cuanto al *civil law*, se evidenció que si bien por ahora la única forma mediante la cual se pueden detectar ultrafalsos es a través de la IA, el solo hecho de emplear esta tecnología generaría una colisión con el ámbito pericial, debido a que las técnicas basadas en sistemas inteligentes transgredirían dos criterios Daubert: la falta de pruebas frente a errores y el déficit de revisión científica. Cabe resaltar que los *deepfakes* también harían irrealizable la constitución de los hechos notorios en

vista principalmente de la era de la posverdad en la que se encuentra el mundo.

Por último, como se ha expuesto anteriormente, las normas procesales, tanto en el sistema anglosajón como en el romano-germánico, no se encontrarían preparadas para dar respuesta a la controversia provocada por los *deepfakes*. Asimismo, los altos costos procesales provenientes de las IA enfocadas en la detección de ultrafalsos no solo representarían un inconveniente para las partes, sino para el proceso de peritaje.

## CONCLUSIONES

A través de la presente investigación se ha llegado a concluir que, si bien la humanidad se ha encontrado viviendo entre noticias falsas primigenias desde que se inventó la imprenta y *deepfakes* primitivos, a partir de la creación de las fotografías, con la injerencia de las tecnologías de medios sintéticos y la distribución de información a grandes escalas por parte de las redes sociales se ha desatado un estado de indefensión para diversos bienes jurídicos, como el honor, la democracia o la verdad.

Si bien el despliegue de sus potencialidades provoca diversas problemáticas, el escenario menos estudiado y quizá el que demanda mayor atención es el del proceso, específicamente el derecho probatorio, ya que se ha podido evidenciar que la nor-

mativa vigente no se encuentra preparada para afrontar los documentos audiovisuales generados por la IA; asimismo, el hecho de que los peritos emplearán IA para detectar IA generaría cuestiones doctrinales y, del mismo modo, se deja en evidencia que los órganos jurisdiccionales no se encontrarían preparados para autenticar este tipo de pruebas.

Finalmente, se hace necesario abordar la problemática *deepfake* de dos formas: primero mediante la generación de políticas que promuevan competencias digitales para la identificación de ultrafalsos por parte de la sociedad; y segundo, la modificación y creación de normas consistentes para dar respuesta a las cuestiones sintéticas.

## REFERENCIAS

- Altuncu, E., Franqueira, V. N. L. y Li, S. (2022). Deepfake: Definitions, Performance Metrics and Standards, Datasets and Benchmarks, and a Meta-Review. *arXiv:2208.10913*. <https://doi.org/10.48550/arXiv.2208.10913>
- Bañuelos, J. (2020). *Deepfake: la imagen en tiempos de la posverdad*. *Revista Panamericana de Comunicación*, 2(1), 51-61. <https://doi.org/10.21555/rpc.v0i1.2315>
- Bravo, A. (2017). Image Manipulation Practices through the History and Evolution of Photography. En J. Jaramillo, A. Burbano, F. Londoño y G. Mejía. (Eds.), *Proceedings of the 23rd International Symposium on Electronic*

- Art ISEA2017 Manizales* (pp. 205-210). Editorial de la Universidad de Caldas.
- Bregler, C., Covell, M. y Slaney, M. (1997). Video Rewrite: Driving Visual Speech with Audio. *SIGGRAPH '97: Proceedings of the 24th Annual Conference on Computer Graphics and Interactive Techniques*, 353-360. <https://doi.org/10.1145/258734.258880>
- Camacho, A. (1998). *Manual de derecho probatorio*. Editorial Themis.
- Cano, S. (2024, 4-5 de abril). La manipulación de la prueba documental a través de la IA: un reto para los jueces [Conferencia]. XIII Fórum Internacional de Expertos y Jóvenes Investigadores en Derecho y Nuevas Tecnologías; Salamanca, España.
- Colegio Técnico del Área de Milwaukee. (2024, 9 de abril). Fake News / AI Deepfakes. <https://guides.matc.edu/fakenews/deepfakes>
- Cootes, T., Edwards, G. y Taylor, C. (2001). Active Appearance Models. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 23(6), 681-685. <https://doi.org/10.1109/34.927467>
- De Ruiter, A. (2021). The Distinct Wrong of Deepfakes. *Philosophy & Technology*, 34, 1311-1332. <https://doi.org/10.1007/s13347-021-00459-2>
- Delfino, R. (2023). Deepfakes on Trial: A Call to Expand the Trial Judge's Gatekeeping Role to Protect Legal Proceedings from Technological Fakery. *uc Law Journal*, 74(2), 293-348.
- Devis Echandía, H. (1984). *Compendio de la prueba judicial*. Rubinzal – Culzoni Editores.
- Eberl, A., Kühn, J. y Wolbring, T. (2022). Using Deepfakes for Experiments in the Social Sciences – A Pilot Study. *Frontiers in Sociology*, 7. <https://doi.org/10.3389/fsoc.2022.907199>
- Gil, R., Virgili-Gomà, J., López-Gil, J.-M. y García, R. (2023). Deepfakes: Evolution and Trends. *Soft Computing: A Fusion of Foundations, Methodologies and Applications*, 27, 11295-11318. <https://doi.org/10.1007/s00500-023-08605-y>
- GOV.UK. (2024). Government Cracks Down on 'Deepfakes' Creation. <https://acortar.link/Y60RFA>
- Jacobsen, B. N. (2023). The Tensions of Deepfakes. *Information, Communication & Society*, 1095-1109. <https://doi.org/10.1080/1369118X.2023.2234980>
- Keith, R. (2024, 30 de enero). Intimate Images Protection Act. *Harper Grey LLP*. <https://acortar.link/1r8x3J>
- Lee, S. (2024, 13 de mayo). AI and Elections: Lessons from South Korea. *The Diplomat*. <https://thediplomat.com/2024/05/ai-and-elections-lessons-from-south-korea/>
- Lubbe, W., Ham-Baloyi, W. y Smit, K. (2020). The Integrative Literature Review as a Research Method: A Demonstration Review of Research on Neurodevelopmental Supportive Care in Preterm Infants. *Journal of Neonatal Nursing*, 26(6), 308-315. <https://doi.org/10.1016/j.jnn.2020.04.006>
- McIntyre, L. (2018). *Post-Truth*. Editorial del Instituto de Tecnología de Massachusetts (MIT).

- Molina, A. C. y Berenguel, O. L. (2022). Deepfake: A Evolução Das Fake News. *Research, Society and Development*, 11(6). <https://doi.org/10.33448/rsd-v11i6.29533>
- Mustak, M., Salminen, J., Mäntymäki, M., Rahman, A. y Dwivedi, Y. K. (2023). Deepfakes: Deceptions, Mitigations, and Opportunities. *Journal of Business Research*, 154. <https://doi.org/10.1016/j.jbusres.2022.113368>
- Naitali, A., Ridouani, M., Salahdine, F. y Kaabouch, N. (2023). Deepfake Attacks: Generation, Detection, Datasets, Challenges, and Research Directions. *Computers*, 12(10). <https://doi.org/10.3390/computers12100216>
- Nieva Fenoll, J. (2018). Repensando Daubert: la paradoja de la prueba pericial. *Civil Procedure Review*, 9(1), 11-26.
- Oxford. (2024). Deepfake. <https://www.oxfordlearnersdictionaries.com/definition/english/deepfake>
- . (2024). Post-Truth. <https://www.oxfordlearnersdictionaries.com/definition/english/post-truth?q=Post-Truth>
- Rare Historical Photos. (2021). How Stalin's Propaganda Machine Erased People from Photographs, 1922-1953. <https://rarehistoricalphotos.com/stalin-photo-manipulation-1922-1953/>
- Rodny-Gumede, Y. (2018). Fake It till You Make It: The Role, Impact and Consequences of Fake News. En B. Mutsvairo y B. Karam. (Eds.), *Perspectives on Political Communication in Africa* (pp. 203-219). Editorial de Palgrave Macmillan.
- Russell, C. (2005). An Overview of the Integrative Research Review. *Prog Transplant*, 15(1). <https://doi.org/10.1177/152692480501500102>
- Sharma, J. y Sharma, R. (2017). Analysis of Key Photo Manipulation Cases and their Impact on Photography. *US University Journal of Arts*, 6(1), 88-99.
- Sodji, L. (2023, 26 de octubre). How We Made David Beckham Speak 9 Languages. *Synthesia*. <https://www.synthesia.io/post/david-beckham>
- Somers, M. (2020, 21 de julio). Deepfakes, Explained. MIT Management Sloan School. <https://mitsloan.mit.edu/ideas-made-to-matter/deepfakes-explained>
- Song, D. (2019, 23 de septiembre). A Short History of Deepfakes. Medium. <https://medium.com/@songda/a-short-history-of-deepfakes-604ac7be6016>
- Suwajanakorn, S., Seitz, S. y Kemelmacher-Shlizerman, I. (2017). Synthesizing Obama: Learning Lip Sync from Audio. *ACM Transactions on Graphics*, 36(4), 1-13. <https://doi.org/10.1145/3072959.3073640>
- The New York Times*. (2009). A Brief History of Photo Fakery. [https://archive.nytimes.com/www.nytimes.com/slideshow/2009/08/23/weekinreview/20090823\\_FAKE\\_SS\\_6.html](https://archive.nytimes.com/www.nytimes.com/slideshow/2009/08/23/weekinreview/20090823_FAKE_SS_6.html)
- Thies, J., Zollhöfer, M., Stamminger, M., Theobalt, C. y Nießner, M. (2019). Face2Face: Real-Time Face Capture and Reenactment of RGB Videos. *Communications of the ACM*, 62(1), 96-104. <https://doi.org/10.1145/3292039>

- Thissen, B. (2024, 6 de febrero). La nueva normativa de la UE penalizará los “manuales de pederastia” y los “deepfakes” de abusos a menores. *Euro News*. <https://acortar.link/FAs4BY>
- Torraco, R. (2016). Writing Integrative Literature Reviews: Using the Past and Present to Explore the Future. *Human Resource Development Review*, 15(4). <https://doi.org/10.1177/1534484316671606>
- Tsay-Vogel, M. (2020). Third-Person Effect. *The International Encyclopedia of Media Psychology*. <https://doi.org/10.1002/9781119011071.iemp0130>
- Tucher, A. (2022). *Not Exactly Lying: Fake News and Fake Journalism in American History*. Editorial de la Universidad de Columbia.
- Van Rijmenam, M. (2022). What is Synthetic Media: The Ultimate Guide. *The Digital Speaker*. <https://www.thedigitalspeaker.com/what-is-synthetic-media-ultimate-guide/>
- Wade, N. (2006). It May Look Authentic; Here's How to Tell It Isn't. *The New York Times*. <https://www.nytimes.com/2006/01/24/science/it-may-look-authentic-heres-how-to-tell-it-isnt.html>
- Waldorf, S. (2021). The Man Who Photographed Ghosts. *Centro Getty*. <https://www.getty.edu/news/the-man-who-photographed-ghosts/>
- Wang, L., Zhou, L., Yang, W. y Yu, R. (2022). Deepfakes: A New Threat to Image Fabrication in Scientific Publications? *Patterns*, 3(5). <https://doi.org/10.1016/j.patter.2022.100509>
- Waters, M. (2017). The Great Lengths Taken to Make Abraham Lincoln Look Good in Portraits. *Atlas Obscura*. <https://www.atlas-obscura.com/articles/abraham-lincoln-photos-edited>
- Westerlund, M. (2019). The Emergence of Deepfake Technology: A Review. *Technology Innovation Management Review*, 9(11), 39-52. <http://doi.org/10.22215/timreview/1282>
- Wolofsky, T. (2014). Photographic Art in the Age of Digital Manipulation: Back to the Darkroom with Michael Lynch and Bruno Latour. *Tooth & Claw*, 11, 89-99.
- Zhang, L. (2023, abril 26). China: Provisions on Deep Synthesis Technology Enter Into Effect. *Library of Congress*. <https://acortar.link/bCKCq7>