

¿CÓMO DEBE RESPONDER EL DERECHO CUANDO NO TODO ES LO QUE PARECE?

Deepfakes, distorsión de la realidad y desafíos jurídicos

En el artículo se plantean algunos de los principales riesgos del uso de videos e imágenes ultrafalsas o *deepfakes*. Asimismo, se reflexiona sobre sus implicaciones jurídicas para responder: ¿qué tan preparadas están las legislaciones para hacer frente a estos riesgos?



MICHELLE AZUAJE-PIRELA
Profesora e investigadora especializada en propiedad intelectual y derecho digital

En marzo de 2021 el mundo se sorprendió con el arresto de Raffaella Spone, una madre de Doylestown (Pensilvania), acusada de manipular fotografías y videos para atacar a algunas compañeras del equipo de animadoras de su hija (Victory Vipers). En esos contenidos, se les mostraba en situaciones que podrían provocar su expulsión del equipo por faltas de conducta, pero además implicaban la difusión de material que involucraba supuestos desnudos o semidesnudos de menores de edad, que negaban haber estado en ellos.

Producto de algunas investigaciones, la madre de una de las víctimas denunció a Spone, a la que detuvieron por la difusión de esos contenidos que enviaba en mensajes de texto anónimos en los que, además, las amenazaba. De

ello trascendió que, un video en el que aparece una joven fumando, tenía las características de un *deepfake*. Los *deepfakes* (también llamados “ultrafalsos” o “ultrafalsificaciones” en español) se refieren a imágenes, videos, clips de audio u otros creados mediante aprendizaje profundo (deep learning, un subcampo de la inteligencia artificial o IA) en los que el contenido se altera con resultados hiperrealistas para que parezca que las personas hicieron o dijeron cosas que realmente no sucedieron o nunca dijeron.

El funcionario policial declaró que detuvo a una madre de los suburbios por utilizar esta tecnología avanzada contra una rival de su hija, creando un video falso increíblemente convincente¹. Así, se viralizó la historia de la que llamaron la *Deepfake mom*.

Sin embargo, la fiscalía nunca pudo confirmar que Spone falsificara las imágenes o videos, por lo que se desestimó esa línea de la acusación, pero en marzo de 2022 fue declarada culpable de ciberacoso y otros cargos por el envío de mensajes con amenazas².

Este caso, de gran notoriedad en Estados Unidos, dejó en el aire dos grandes preguntas. Primera: ¿El video era real o ciertamente era un *deepfake*? Algunos expertos consultados ►►

1. Korducki, Kelli M. (2021): “The World Thought This Cheer Mom Created a Deepfake to Harass Her Daughter’s Rival—but the Real Story Is Way More Confusing (and Bizarre)”, *Cosmopolitan*. Disponible en: <https://www.cosmopolitan.com/lifestyle/a37377027/deep-fake-cheer-scandal/>
2. Heffintrayer, K. (2022): “Jury Finds Chalfont Cheer Mom Guilty on All Charges”, *NorthPennNow*. Disponible en: <https://northpennnow.com/jury-finds-chalfont-cheer-mom-guilty-on-all-charges-p5350-176.htm#>



Palabras clave: posverdad, informaciones ultrafalsas, *deepfakes*, desinformación.

How should the law respond when not everything is what it seems?

DEEPPAKES, DISTORTION OF REALITY AND LEGAL CHALLENGES

The article discusses some of the main risks of the use of “ultrafake” or “deepfakes” videos and images. It also reflects on their legal implications in order to answer: how prepared are legislations to face these risks?

Keywords: post-truth, ultra-false information, *deepfakes*, disinformation.

El abaratamiento, sofisticación y masificación de estas tecnologías puede perjudicar a un sinnúmero de individuos y convertirse en un verdadero dolor de cabeza para la justicia

en canales nacionales opinaron que, con la tecnología disponible de forma gratuita en ese entonces, era imposible que Spone (sin conocimientos técnicos) pudiese lograr efectos tan realistas porque no había suficiente data disponible sobre las animadoras para lograrlo, así que dudaron que fueran deepfakes, pero debido a la forma en que se manejó el caso, la duda quedó sembrada.

Y es aquí donde viene la segunda gran pregunta: ¿Qué faltó por hacer? ¿qué se necesitaba para saber si el video era real o no? Se necesitaba, por ejemplo, utilizar sistemas de IA para detectarlo, o bien, personas —como peritos— que apliquen otras técnicas forenses específicas para revisar el contenido, buscando elementos que permitiesen concluir si era real o no. ¿El problema? Incluso en Estados Unidos (un país a la vanguardia del desarrollo tecnológico), solo unos pocos profesionales podrían hacer esas diligencias. Entonces,

¿cómo debe responder el derecho cuando no todo es lo que parece? ¿qué tan preparadas están las legislaciones, los tribunales o una pequeña oficina policial para hacer frente a estos y otros riesgos de los deepfakes?

Desafío legal

Los deepfakes pueden emplearse para poner a personas en situaciones en las que no estuvieron realmente (como el papa Francisco, caminando por las calles usando un abrigo blanco de Balenciaga) sin que necesariamente podamos advertir la falsedad del contenido. Pueden hacerse con fines humorísticos o de parodia (en principio, amparados por la libertad de expresión), o para mejorar algunas obras cinematográficas, por ejemplo, a través de “doblajes sintéticos sincronizados” (como promete TrueSync), pero también con los fines más perturbadores.

Por ejemplo, se han utilizado para “revivir” digitalmente a personas fallecidas como ocurrió con el deepfake de Lola Flores para una campaña publicitaria en España, o con el deepfake del chef Antony Bourdain para el documental Roadrunner en el que se recreó su voz para pronunciar diálogos que nunca dijo ante las cámaras, lo que levantó ciertos cuestionamientos éticos.

En otros contextos se han utilizado deepfakes para campañas electorales (como en India y Corea del Sur), o para difundir fake news. Por esa y otras razones se teme porque se avecina un “infocalipsis”³ que puede incluso afectar la democracia, la seguridad nacional y hasta se teme que se usen nuevas y sofisticadas técnicas de desinformación en conflictos armados internacionales⁴.

En sus orígenes los deepfakes atacaron principalmente a personas famosas (es decir, a unas pocas personas, como Gal Gadot, Taylor Swift, Scarlett Johansson, más recientemente influencers, escritoras, entre otras), pero el abaratamiento, sofisticación y masificación de estas tecnologías puede llegar a perjudicar a un sinnúmero de individuos a nivel mundial y convertirse en un verdadero dolor de cabeza para el sistema de justicia. En algunos casos, los problemas tendrán

solución en la legislación en vigor; en otros, la desbordarán y no siempre encontrarán acomodo. Por lo tanto, se necesitan algunas reflexiones, adecuaciones y resguardos. Y, sobre todo, se necesitan conocimientos porque no se puede legislar o juzgar sobre algo que no se conoce o no se entiende.

Al margen de las “curiosidades” y de los ejemplos de afectaciones a personas de renombre mundial, existe una gran preocupación y es que pueden usarse —y, de hecho, se han usado— para perjudicar a cualquier persona y esto va en aumento. Así, con ellos se pueden encontrar nuevos canales para cometer delitos (por ejemplo, ya se han replicado las voces de empresarios para realizar intimidaciones y estafas⁵).

Si bien los casos de personas no famosas víctimas de videos ultrafalsos todavía no son muy extendidos, ya se conocen ofertas de servicios de pago para poner a cualquiera en contextos pornográficos no consentidos, lo que es solo una muestra de los grandes desafíos jurídicos que tenemos por delante. Derechos como la imagen, el honor y la reputación están en especial riesgo. En ese sentido, la proliferación de deepfakes ha afectado especialmente a mujeres, a las que se ha puesto en contextos pornográficos sin su consentimiento generando

irreparables daños a sus víctimas. Por ello, se están impulsando modificaciones para ampliar las leyes contra la “pornovenganza”.

Por ejemplo, en algunos estados de Estados Unidos —como Virginia— ya se ha actualizado la legislación para imponer penas y multas al respecto. Hay iniciativas que buscan prohibir la creación de deepfakes en algunos contextos (como electorales, en Texas), así como facilitar el desarrollo de herramientas tecnológicas que permitan detectarlos. Otras medidas como la contenida en la propuesta del Reglamento Europeo sobre Inteligencia Artificial plantean la obligación de informar de que dicho contenido se ha generado por medios automatizados, salvo excepciones que respondan a fines legítimos (por ejemplo, por la aplicación de la ley o por el ejercicio de la libertad de expresión).

Reflexión final

Los recientes desarrollos tecnológicos nos están llevando a transitar por pasajes inimaginables que muchos no ven y otros solo ven como una “curiosidad”, pero que encierran debates profundos y que han afectado y pueden seguir afectando los derechos de miles de personas a nivel mundial. Ello impone la

obligación de ser proactivos y reflexionar sobre los retos globales que hoy enfrentamos y, muy especialmente, de preguntarnos qué puede o debe hacer el derecho al respecto. Sabemos que deben tomarse algunas medidas, sí, pero... ¿lo estamos haciendo?, ¿saben o deben saber los operadores de justicia —incluso en una pequeña localidad— lo que es un deepfake?, ¿lo saben los abogados?, ¿deberían saberlo?, ¿quién nos prepara para ello?, ¿cómo se entendería correctamente probada la falsificación de uno de estos videos, audios o imágenes en juicio?, ¿tenemos las herramientas para ello?

Bibliografía

Schick, N. (2020): *Deepfakes. The coming infocalypse*. New York, Twelve.
Byman, Daniel L.; Gao, C.; Meserole, C. y Subrahmanian, V.S. (2023): *Deepfakes and International conflict*. The Brookings Institution.
Bedecarratz Scholz, F. y Navarro Dolmestch, R. (2022): “Medidas de ciberprotección para los creadores culturales”, en Azuaje-Pirela, M. y Forero, F. A. (coord.): *Tecnologías emergentes. Qué son y cómo aprovecharlas en las industrias creativas y culturales*, pp. 87-106.



3. Schick, N. (2020): *Deepfakes. The coming infocalypse*. New York, Twelve.

4. Byman, D. L.; Gao, C., Meserole C. & Subrahmanian, V. S. (2023): “Deepfakes and International conflict”, *Foreign Policy* at Brookings. Disponible en: https://www.brookings.edu/wp-content/uploads/2023/01/FP_20230105_deepfakes_international_conflict.pdf

5. Bedecarratz Scholz, F. & Navarro Dolmestch, R. (2022): “Medidas de ciberprotección para los creadores culturales”, en Azuaje-Pirela, M. & Forero, F. A. (coord.): *Tecnologías emergentes. Qué son y cómo aprovecharlas en las industrias creativas y culturales*, págs. 87-106, CERLALC. Disponible en: <https://cerlalc.org/publicaciones/tecnologias-emergentes-que-son-y-como-aprovecharlas-en-las-industrias-creativas-y-culturales/>