



Menos del 1% de las personas puede detectar falsificaciones generadas por IA, según un estudio

El trabajo de la multinacional iProov revela poca conciencia de los usuarios ante los deepfakes, videos e imágenes hechas con IA.

Ignacio Arriagada M.

En plena era digital, la manipulación de imágenes, audios y videos ha alcanzado un nivel de sofisticación y realismo sin precedentes gracias al uso de técnicas de inteligencia artificial (IA). Este tipo de contenido, denominado mundialmente como deepfake, muestra algo que, a la vista y oído de cualquier persona, parece verdadero y real. Sin embargo, es totalmente falso.

A medida que los deepfakes se vuelven más auténticos y su viralización aumenta por redes sociales, se ha convertido en un reto detectar si ha sido creado por la IA. Así lo confirma un estudio realizado por iProov, proveedor mundial de soluciones científicas para la verificación biométrica de identidad, que indica que menos del 1% de las personas pudo distinguir con precisión el contenido real del falso en todos los estímulos, que incluían imágenes y videos.

RESULTADOS

El estudio contó con la participación de 2.000 habitantes de Estados Unidos y del Reino Unido, de entre los 18 y 65 años y más, a quienes se les expuso una serie de contenidos reales y deepfakes. Los principales hallazgos fueron los siguientes.

-La detección de deepfakes falla: solo 0,1% de los encuestados logró identificar correctamente todos los estímulos deepfake y reales en este estudio, en el que se preparó a los participantes para que buscaran el contenido falso.

-Generaciones mayores, las más vulnerables: el trabajo encontró que el 30% de las personas de 55-64 años y el 39% de los mayores de 65 años nunca habían oído hablar de los deepfakes, lo que destaca una importante brecha de conocimiento y una mayor susceptibilidad en este grupo.

-Videos, los más preocupantes: los videos son particularmente desafiantes, ya que solo el 9% de los participantes pudo identificar todos

los deepfakes. Esta vulnerabilidad genera serias preocupaciones sobre la posibilidad de fraude basado en videos, como suplantación de identidad en videollamadas o en escenarios en los que se utiliza la verificación de video para verificar la identidad.

-Los deepfakes están por todas partes, pero no se comprenden: si bien la preocupación por el contenido audiovisual falso está aumentando, muchos desconocen esta tecnología. Uno de cada cinco de los encuestados (22%) nunca había oído hablar de esto antes del estudio.

-El exceso de confianza está muy extendido: a pesar de su bajo rendimiento, los consultados mantuvieron una confianza excesiva en sus habilidades para detectar deepfakes (más del 60%), independientemente de si sus respuestas eran correctas o no. Esto fue especialmente así en el caso de los adultos jóvenes (de 18 a 34 años).

-La confianza se ve afectada: las plataformas de redes sociales se consideran caldo de cultivo para deepfakes. De



Las habilidades para detectar un contenido audiovisual adulterado es nulo para algunos internautas.



2.000 PERSONAS de EE.UU. y del Reino Unido participaron en la investigación de iProov.

EL 30% DE LOS participantes de 55-64 años reconoció nunca haber oído hablar de los deepfakes.

hecho, Meta (49%) y TikTok (47%) son los lugares donde más se encuentran deepfakes en línea, según advierte iProov. Esto, a su vez, ha redu-

cido la confianza en la información y los medios en línea. En ese contexto, el 49% de los entrevistados confía menos en las RR.SS. tras enterarse de la existencia de deepfakes y solo uno de cada cinco de ellos denunciaría una sospecha en esas plataformas.

-Se necesitan mejores mecanismos de concientización y denuncia: menos de un tercio de las personas entrevistadas (29%) no toman ninguna medida cuando se encuentran con una presunta falsificación profunda, lo

que probablemente se debe a que el 48% de ellos dice que no sabe cómo denunciar, mientras que a una cuarta parte no le importa si ve una presunta deepfake.

-La mayoría de los consumidores no verifican activamente la autenticidad de la información en línea: a pesar de la creciente amenaza de la desinformación, solo uno de cada cuatro de los participantes busca fuentes de información alternativas si sospecha que se trata de una falsificación profunda.