

Un estudio reveló que la mayoría de las personas no distingue las voces clonadas con IA de las humanas

04/10/2025



La inteligencia artificial avanza a pasos acelerados y una de sus áreas que más creció en los últimos meses es la clonación de voz. Lo que hace apenas unos años parecía ciencia ficción, hoy ya es una realidad: investigadores británicos comprobaron que las voces creadas con software de IA suenan tan reales que la mayoría de las personas no logra diferenciarlas de una voz humana.

El estudio, publicado en la revista científica PLOS One y realizado por especialistas de la Queen Mary University of London, no solo expuso hasta qué punto la tecnología de síntesis de voz se volvió naturalista, sino también la facilidad con la que se pueden replicar las voces de las personas.

Los científicos encargados de la investigación compararon grabaciones de personas reales con **audios** generados por dos sistemas distintos: uno capaz de **clonar voces específicas a partir de grabaciones cortas** y otro que utiliza **modelos de voz generales** para crear locuciones que no corresponden a ningún individuo en particular.



Con solo unos minutos de audio y programas gratuitos se pueden crear deepfakes de voz que plantean un debate urgente: cómo convivir con esta nueva era tecnológica sin poner en riesgo la confianza social. (Foto: Adobe Stock)

Los participantes del experimento debían evaluar cuáles voces sonaban más realistas, dominantes o confiables. El resultado fue contundente: **los audios sintéticos fueron prácticamente indistinguibles de los humanos.**

Clonar una voz ya es fácil, barato y peligroso

Uno de los puntos que más preocupó a los investigadores es la **facilidad con la que se pueden generar estas voces falsas.** Según explicó Nadine Lavan, psicóloga y coautora del estudio, el proceso requirió una mínima experiencia, apenas unos

minutos de grabaciones de voz y **casi ningún costo económico**, ya que se utilizó software comercial disponible para cualquiera.

Este nivel de accesibilidad despierta alertas en el plano de la **seguridad y la confianza digital**. Las voces clonadas podrían usarse en **estafas telefónicas**, suplantación de identidad, **manipulación política** o campañas de desinformación, ya que no solo suenan auténticas, sino que también transmiten sensaciones de credibilidad y cercanía.

¿Qué son los deepfakes de voz?

Los **deepfakes de voz** son audios generados por inteligencia artificial que imitan con **gran precisión** la forma de hablar, el tono y las características vocales de una persona real. Se logran con apenas unos minutos de grabaciones de una persona o **programas y clips disponibles en internet** y redes sociales, como *reels*, *stories* o videos de YouTube.

¿Por qué son un riesgo?

- Pueden usarse para estafas telefónicas en las que un delincuente se hace pasar por un familiar o jefe y **pedirle dinero o que brinde información confidencial**
- Sirven para suplantar la identidad de figuras públicas o **manipular** declaraciones, especialmente en **campañas electorales**.
- Dificultan la verificación de la información, ya que la voz siempre fue un **factor de confianza** en la comunicación.

El desafío que viene

Para los especialistas, los resultados de la investigación

marcan un punto de inflexión. “Era solo cuestión de tiempo hasta que la tecnología comenzara a producir un habla naturalista y humana”, advirtió Lavan, quien subrayó la necesidad de **entender cómo las personas perciben estas voces** y desarrollar estrategias para detectar fraudes antes de que se masifiquen.

La clonación de voz con IA es apenas una de las aristas del universo de los **deepfakes**, que también incluye imágenes y **videos** hiperrealistas. Ahora, con audios imposibles de distinguir del original, los expertos plantean un debate urgente: **cómo convivir con esta nueva tecnología sin poner en riesgo la confianza social.**

Fuente: TN