

# La nueva IA de Microsoft hace rapear a la Mona Lisa

La Mona Lisa puede hacer algo más que sonreír gracias a la nueva tecnología de inteligencia artificial de Microsoft.

La semana pasada, investigadores de Microsoft presentaron un nuevo modelo de inteligencia artificial que puede tomar una imagen fija de una cara y un clip de audio de alguien hablando y crear automáticamente un video realista de esa persona hablando. Los videos -que pueden hacerse a partir de rostros fotorrealistas, dibujos animados o ilustraciones- se completan con una convincente sincronización labial y movimientos naturales de la cara y la cabeza.

En un video de demostración, los investigadores mostraron cómo animaban a la Mona Lisa para que recitara un rap cómico de la actriz Anne Hathaway.

Los resultados del modelo de IA, llamado Vasa-1, son tan divertidos como un poco chocantes por su realismo. Según Microsoft, esta tecnología podría utilizarse en educación o para “mejorar la accesibilidad de personas con problemas de comunicación”, o incluso para crear compañeros virtuales para humanos. Pero también es fácil ver cómo se podría abusar de la herramienta y utilizarla para suplantar a personas reales.

Es una preocupación que va más allá de Microsoft: a medida que surgen más herramientas para crear imágenes, videos y audio convincentes generados por IA, a los expertos les preocupa que su uso indebido pueda dar lugar a nuevas formas de desinformación. A algunos también les preocupa que la tecnología pueda perturbar aún más las industrias creativas, desde el cine hasta la publicidad.

De momento, Microsoft no tiene previsto hacer público el modelo Vasa-1 de forma inmediata. La medida es similar a la forma en que OpenAI, socio de Microsoft, está gestionando las preocupaciones en torno a su herramienta de video generado por IA, Sora.

OpenAI presentó Sora en febrero, pero hasta ahora solo la ha puesto a disposición de algunos usuarios profesionales y profesores de ciberseguridad con fines de prueba.

“Nos oponemos a cualquier comportamiento para crear contenidos engañosos o dañinos de personas reales”, afirmaron investigadores de Microsoft en una entrada de su blog. Pero, añadieron, la empresa “no tiene planes de lanzar” el producto

públicamente “hasta que estemos seguros de que la tecnología se utilizará de forma responsable y de acuerdo con la normativa adecuada”.

El nuevo modelo de inteligencia artificial de Microsoft se entrenó con numerosos videos de caras de personas hablando, y está diseñado para reconocer los movimientos naturales de la cara y la cabeza, incluidos “el movimiento de los labios, la expresión (no labial), la mirada y el parpadeo, entre otros”, explicaron los investigadores. El resultado es un video más realista cuando Vasa-1 anima una foto fija.

Por ejemplo, en un video de demostración con un clip de alguien que parece agitado, aparentemente mientras juega a videojuegos, la cara que habla tiene las cejas fruncidas y los labios fruncidos.

Si se observan de cerca, siguen apareciendo indicios de que los videos han sido generados por una máquina, como parpadeos poco frecuentes y movimientos exagerados de las cejas. Pero Microsoft cree que su modelo “supera con creces” a otras herramientas similares y “allana el camino para interactuar en tiempo real con avatares realistas que emulan comportamientos conversacionales humanos”.

Con información de CNN en español