

Si parpadean, se lo van a perder

Emilio J. Molina

ARP-Sociedad para el Avance del Pensamiento Crítico

Sobre inteligencia artificial y desinformación

Hace casi exactamente un año charlábamos en *Escépticos en el Pub Madrid*¹ sobre los (entonces) emergentes *deep fakes*, y cómo iban a suponer un revulsivo en la cantidad de desinformación a la que estaremos expuestos. El núcleo de la charla era, básicamente, que la manipulación audiovisual de la información no es nada nuevo y que no hay más que comprobar las aparentemente inevitables cadenas de *Whatsapp* para comprobar que ni siquiera es necesaria sofisticación alguna para que los bulos se difundan a sus anchas. Sin embargo, los nuevos avances en inteligencia artificial (IA) vienen a subir la apuesta sobre el problema, hipersofisticando la calidad y persuasión del engaño.

En un alarde de dotes clarividentes, comentaba yo en la charla que los que tenían en ese momento las herramientas más avanzadas no las habían liberado

aún al público, en un ejercicio de autocontrol y ética para no desatar a la bestia sin tener claro cómo (o si se podía) domesticarla. Creo que fue cuestión de dos semanas tras la charla que algunos de estos modelos pasaran a formar parte del código publicado en repositorios al alcance de todo el mundo.

En ese momento, hace un año, insisto, habíamos pasado a su vez desde algo que nos parecía imposible (que un ordenador nos generara una imagen realista por su cuenta) a algo que, más o menos rudimentario, ya llegaba al nivel de hacernos dudar (y, en ocasiones, a colárnosla).

Pero esto es el mundo de la informática y, como muchos saben, aquí funcionan a menudo las leyes de evolución cuadrática. Los avances van siendo exponenciales, y donde el año pasado aún teníamos trucos de *Blade Runner* tan toscos como fijarnos en cuántos

Es fácil ser escéptico de aquello que nos suena mal, pero muy difícil que no nos cuelen aquello que ya estamos predispuestos a creer. Subid el volumen de la desconfianza ante esto último

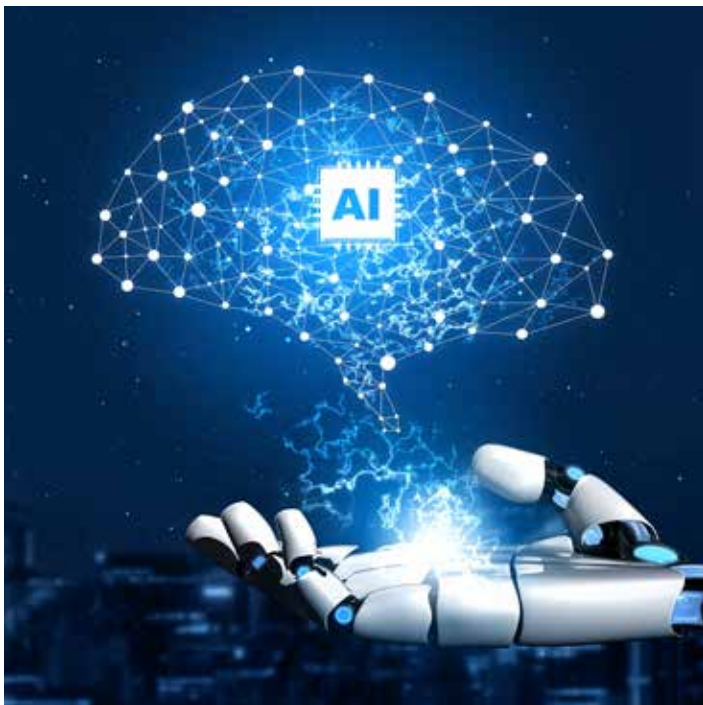


Imagen de Tung Nguyen en Pixabay

dedos les ponían estas IA a las manos, este año necesitamos recurrir a asuntos muchísimo más sofisticados (en la charla se desgranaban algunos). Donde el año pasado el problema era más a nivel de imágenes fijas o de vídeos no muy complejos, este año tenemos hiperresoluciones animadas prácticamente perfectas. Pero aún hay más. Mucho más.

Durante agosto, Álex², un colega de profesión que se encarga de los fantásticos *Deep Fakes* del programa de televisión *El Intermedio*, quería probar en el grupo de chat de Telegram de *El Mundo Today* una nueva tecnología de conversión de voz, en la que le pasas un audio cualquiera a una IA alimentada con los audios de la persona a la que quieres imitar, y te devuelve tu mismo audio con su timbre. En este grupo, Kike García, uno de los creadores de EMT, lleva un tiempo enviando notas de voz de unos diez minutos a modo de parodia de los típicos audios insulsos que muchos recibimos a veces, a los que ha terminado dando forma de monólogos. Aprovechando sus vacaciones (las del propio Álex y las de Kike), nos pidió a los miembros del canal que mandáramos audios breves, que él usaría para intentar transformar en la voz de Kike. Incluso siendo una tecnología reciente que mi colega aún estaba tratando de domesticar, los resultados fueron apabullantes. A poco que la persona copiara sus modismos (y no tuviera acentos muy reconocibles, que ahí se veían fácilmente las costuras), los audios tenían una similitud pasmosa. Y hablamos de una voz que, por lo comentado (y por sus divertidos contenidos audiovisuales), estamos acostumbrados a oír.

En la charla ya comenté que, probablemente, los *deep fakes* de audio vayan a convertirse en una herramienta mucho más peligrosa que los de vídeo, por el hecho de que en vídeo aún hay muchos paráme-

tros que nos pueden dar una pista de que algo falla (dientes moviéndose, sombras no exactamente en su sitio, ciertos desenfoces...), mientras que en un audio contamos con muchas menos pistas. Y parte del problema también es que estas modificaciones, si no se dan ya, están a un paso de poderse hacer en tiempo real. Imaginad qué significa esto para asuntos como el *phishing* y otros tipos de estafas que actualmente se hacen vía texto con mucha menor sofisticación (el día que los ciberdelincuentes aprendan a escribir sin faltas de ortografía, el mundo será un lugar mucho más inseguro; al menos, para quienes sepan detectarlas).

Por si fuera poco, la aparición de herramientas como ChatGPT ha venido a lanzar un órdago muchísimo mayor si cabe. Pero este es un tema tan inmenso que excede de lo que quería tratar en el artículo, más allá de que se puede usar para generar desinformación de forma muy rápida y con un grado de calidad superior a la media.

La charla de *Escépticos en el Pub* se grabó con una calidad de sonido bastante pobre, por desgracia. Pero apenas unos meses tras la charla, también aparecían herramientas como *Whisper*³, de la misma OpenAI que liberó algunos de los modelos comentados, y que permite una transcripción a texto de bastante calidad. Espero que cuando lean estas líneas ya haya sacado tiempo para añadir subtítulos a la charla y que podáis comprobar cómo ha envejecido en solo un año (por ejemplo, con los recientes vídeos virales de traducciones a otros idiomas con la misma voz del original).

No es casualidad que haya algunos productos audiovisuales como el recomendable *podcast Titania*⁴ y la espectacular *Mission Impossible: Sentencia Mortal (Parte 1)*⁵, donde este tipo de avances, mucho más cercanos de lo que nos sentiremos cómodos de reconocer, forman parte fundamental de la trama.

Os dejo dos consejos de gurú clarividente (*disclaimer*: véase el segundo párrafo):

1. No gastéis dinero en cursos de *prompt engineer* ni cosas así; está avanzando todo a tal velocidad que la semana que viene podrían no tener ya sentido.
2. Es fácil ser escéptico de aquello que nos suena mal, pero muy difícil que no nos cuelen aquello que ya estamos predisuestos a creer. Subid el volumen de la desconfianza ante esto último, máxime cuando vuestros hijos os pidan dinero porque les han robado el móvil o un príncipe nigeriano os llame en persona para nosequé.

Notas:

1. <https://www.escepticos.es/node/8563>
2. <https://twitter.com/thehardmenpath>
3. <https://openai.com/research/whisper>
4. https://www.ivoox.com/podcast-titania_sq_f11946762_1.html
5. <https://www.filmaffinity.com/es/film945258.html>