INTERNET | FALSOS CONTENIDOS

**Sitios web falsos   
generados por IA que generan desinformación**

*Se están generando todo tipo de granjas de clics automatizados con la única intención de generar ingresos por publicidad, con información poco confiable.*

**Por Julián Marcelo Zappia.  
Diplomatic Week.**

Una de las principales preocupaciones que ha llevado a los expertos en IA como Geoffrey Hinton a temer esta tecnología es el mal uso que algunos le pueden dar a la misma.

En España se han visto ejemplos como los deepfakes, vídeos e imágenes hechos con inteligencia artificial que usualmente se usan con motivos políticos.

Según informa el sitio de calificación de webs NewsGuard, los usuarios están dándole un nuevo uso a la inteligencia artificial: crear sitios y publicaciones en webs falsas para generar las clásicas granjas de contenido.

En definitiva, son sitios de spam. La idea es que, gracias a la atracción en Internet, estos sitios puedan generar ingresos publicitarios de clics de usuarios. Esto es algo que ya se estaba haciendo pero con un proceso muchísimo más manual y laborioso, y que ha llevado a muchos a usar IA para crear estas webs.

Y es que la adopción de herramientas como ChatGPT ha abierto las puertas a la creación de estos sitios.

NewsGuard explica que se han identificado hasta 49 páginas que parecen "estar escritas casi en su totalidad por software de inteligencia artificial". Webs que intentan pasar desapercibidas para que no sean objetivo de los controles de Google pero que pueden generar una gran cantidad de contenido de forma automática, llegando a cientos de publicaciones al día.

**Sitios de spam con IA.**

Según NewsGuard, esta clase de sitios "no revelan la propiedad o el control" de sus contenidos por motivos evidentes. La temática de las publicaciones da exactamente igual; tratan temas de política, salud, entretenimiento, finanzas y tecnología.

Lógicamente muchos de estos contenidos provienen de bases de datos que recogen artículos ya escritos de medios prestigiosos, lo que resulta en traducciones poco elaboradas.

Salvando esas reproducciones, NewsGuard detalla que buena parte del contenido publicado es completamente falso. Casi todo el contenido presenta un lenguaje suave y frases repetitivas, características de la inteligencia artificial. A eso se le suman nombres genéricos que están cargados de publicidad.

La mayoría son reproducciones o resúmenes de artículos de medios internacionales, como el The New York Times o la BBC.

El que mucho de este contenido no sea falso tiene dos explicaciones. La primera es que son reproducciones de historias verdaderas con investigación detrás y la segunda es que el contenido directamente falso tarda menos en ser eliminado en caso de ser detectado. No obstante, esto no ha evitado que estas webs proliferen.

Un ejemplo es la web ya extinta CelebritiesDeaths, que recogía una noticia falsa sobre cómo Joe Biden, actual presidente de los Estados Unidos, había fallecido, según informa el sitio The Verge.

Este caso es especialmente flagrante, ya que los párrafos siguientes incluían directamente un mensaje de error del chatbot que generó dicho mensaje, y los responsables del sitio no dudaron en publicarlo. El mensaje de error afirmaba que no era ético fabricar noticias sobre la muerte de alguien.

Además de los errores de traducción y gramática existente en estos artículos, existen herramientas directas que permiten verificar cuándo un texto es generado por IA. Es el caso de GPTZero, que sobre el papel es capaz de detectar texto generado por un chatbot, y sirvió para que NewsGuard pudiera hallar muchos de los problemas de estas webs.

El aumento de la accesibilidad en estos chatbots y sobre todo la popularización de las herramientas de IA generativa podrían hacer que las predicciones de muchos de estos expertos se hagan realidad: una era de desinformación en la que el usuario no pueda verificar qué es real y qué no. Una expansión de estos problemas podría obligar a los gobiernos de todo el mundo a tomar medidas para regular el uso de estas tecnologías.