Demuestran científicos que modelo de inteligencia artificial informa y engaña mejor

Image not found or type unknown

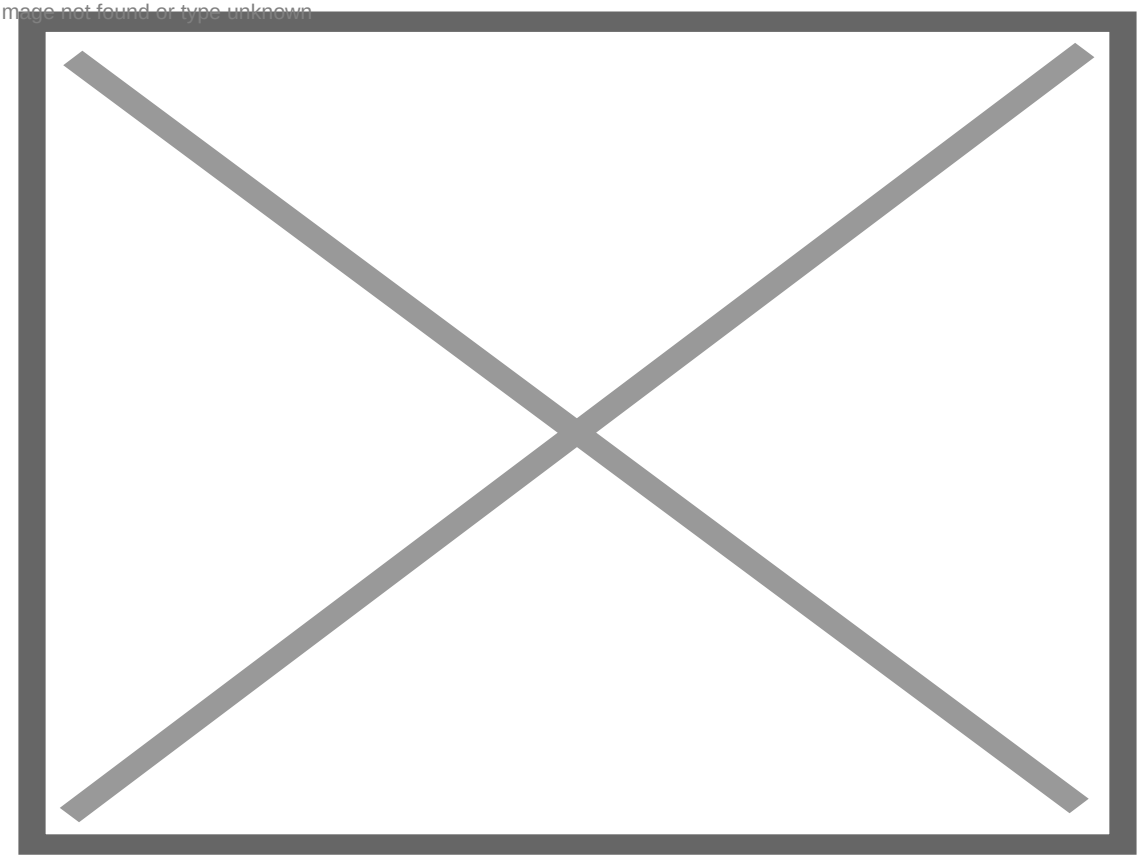

Imagen: Prensa Latina.

Washington, 28 jun (RHC) Investigadores de la Universidad de Zúrich, Suiza, demostraron el poder de la inteligencia artificial -IA-, para informar y engañar, lo que plantea preguntas críticas sobre el futuro de ecosistemas de información, divulgó hoy Science Advances.

El estudio se centró específicamente en GPT-3, un modelo de lenguaje de IA avanzado de OpenAI, para determinar sus riesgos y beneficios potenciales en la generación y difusión de (des)información.

Con 697 participantes la investigación buscó evaluar si las personas podían diferenciar entre desinformación e información precisa presentada en forma de tweets, y si discernían quién lo había escrito, si un usuario genuino de la red o la tecnología digital.

Los temas tratados incluyeron el cambio climático, la seguridad de las vacunas, la pandemia de la COVID-19, la teoría de la tierra plana y los tratamientos homeopáticos para el cáncer.

GPT-3 demostró la capacidad de generar información precisa y, en comparación con los tweets de usuarios reales de Twitter, información más fácilmente comprensible.

Sin embargo, los investigadores también descubrieron que el modelo de lenguaje de IA tenía una habilidad inquietante para producir desinformación altamente persuasiva.

En un giro preocupante, los participantes no pudieron diferenciar de manera confiable entre los tweets creados por GPT-3 y los escritos por usuarios genuinos.

“Nuestro estudio revela el poder de la IA para informar y engañar, lo que plantea preguntas críticas sobre el futuro de los ecosistemas de información”, dijo Federico Germani, uno de los investigadores.

De acuerdo con los científicos, estos hallazgos sugieren que las campañas de información creadas por GPT-3, basadas en indicaciones bien estructuradas y evaluadas por humanos capacitados, serían más efectivas, por ejemplo, en una crisis de salud pública que requiere una comunicación rápida y clara al público.

También plantean preocupaciones importantes con respecto a la amenaza de que la IA perpetúe la desinformación, particularmente en el contexto de la difusión rápida y generalizada de información errónea y desinformación durante una crisis o evento de salud pública.

El estudio revela que los sistemas impulsados por IA podrían explotarse para generar campañas de desinformación a gran escala sobre potencialmente cualquier tema, poniendo en peligro no solo la salud pública sino también la integridad de los ecosistemas de información vitales para el funcionamiento de las democracias.

En este contexto, los expertos hicieron un llamado a los legisladores para que respondan con regulaciones estrictas, basadas en evidencia y éticamente informadas para abordar las amenazas potenciales que plantean estas tecnologías disruptivas y garantizar el uso responsable de la IA. (Fuente: Prensa Latina).

---

https://www.radiohc.cu/noticias/ciencias/326871-demuestran-cientificos-que-modelo-de-inteligencia-artificial-informa-y-engana-mejor

RADIO HABANA CUBA

Radio Habana Cuba