



¿Grok, es esto real?

Análisis crítico del fact-checking en X y su espacio en la discusión pública

Rebecca Marilyn Rodríguez Ríos, Consuelo Lemus Pool

UNIVERSIDAD AUTÓNOMA DE TAMAULIPAS

Resumen. Analizamos el papel de Grok, inteligencia artificial generativa integrada en la red sociodigital X, como símbolo del tránsito de la verificación de hechos desde el terreno humano hacia el algorítmico. Su origen, ligado a la visión de Elon Musk y a la idea de un agente guardián de la verdad, plantea dilemas éticos y epistemológicos sobre el uso de sistemas automatizados en la lucha contra la desinformación. El objetivo es reflexionar críticamente sobre los riesgos, límites y responsabilidades que surgen al delegar la noción de verdad a entidades carentes de intencionalidad moral o comprensión contextual. A partir del análisis del fact-checking algorítmico, concluimos que Grok representa una solución aparente frente a la saturación informativa; pero, al mismo tiempo, genera nuevas zonas de ambigüedad epistemológica y evidencia la paradoja contemporánea: cuanto más se busca la certeza, más se confía en entidades incapaces de experimentarla.

Abstract. We analyze the role of Grok, a generative artificial intelligence integrated into the social network X, as a symbol of the transition of fact-checking from the human to the algorithmic realm. Its origin, linked to Elon Musk's vision and the idea of an AI "guardian of truth," raises ethical and epistemological dilemmas regarding the use of automated systems in the fight against disinformation. The objective is to critically reflect on the risks, limits, and responsibilities that arise when delegating the notion of truth to entities lacking moral intentionality or contextual understanding. Based on an analysis of algorithmic fact-checking, we infer that Grok represents an apparent solution to informational overload while simultaneously generating new areas of epistemological ambiguity, revealing a contemporary paradox: the more certainty we seek, the more we rely on entities incapable of experiencing truth.

Grok: el no-ser de la verdad

En X (anteriormente Twitter) es cada vez más común presenciar la intervención de una inteligencia artificial generativa (IAG) en las discusiones que surgen a menudo de las tendencias que el mismo algoritmo selecciona para el feed y para arrojar a las fauces de usuarios que desean entrar a un nuevo debate u opinar en el "hilo" de un storytelling. Este

no-ser algorítmico, bautizado como Grok, emerge como una herramienta multifuncional y su fin más controvertido es el principal interés de este análisis: su herramienta de fact-checking dentro de la plataforma.

La irrupción de Grok es especialmente significativa si consideramos su origen. Inicialmente, Elon Musk presentó su visión públicamente según reportó Bailee Hill en el portal de Fox News el 17 de abril de 2023. Su enfoque nació directamente de la preocupación, a decir del empresario, por la falta de regulación y el potencial dañino de modelos de IAG como ChatGPT a la civilización. Musk, las IAG no reguladas potencialmente podrían ser utilizadas para “atacar la democracia a través del control de las mentes de los americanos”. En esta entrevista, Musk introduce a TruthGPT como:

Una IAG cuyas características, a decir del empresario, se dedican a la búsqueda de la verdad máxima que intenta comprender la naturaleza del universo. Y creo que esta podría ser la mejor vía hacia la seguridad, en el sentido de que es poco probable que una IA que se preocupa por comprender el universo aniquile a los seres humanos, ya que somos una parte interesante del universo.

TruthGPT nunca salió a la luz y la idea inicial nunca se materializó bajo esa entidad idealizada. Sin embargo, este discurso dio pie al nacimiento de Grok el 4 de noviembre de 2023. De acuerdo con la descripción que aparece en el centro de ayuda de X, es “un asistente de inteligencia artificial que ayuda a completar tareas, como responder preguntas, resolver problemas y realizar lluvias de ideas” (X Corp, 2025). Además, hay una particularidad clave: su conexión directa y en tiempo real con la vasta y turbulenta plataforma X.

La diferencia entre la promesa de TruthGPT, que sería guardiana de la verdad máxima, y Grok, ese ente algorítmico con un humor algo picante y juguetón, dependiendo del tono con el que se la pida participar,, que se une a las discusiones para dotar de “razón” al usuario que lo convoca a través del @ y unas palabras mágicas (“¿Grok, es esto real?”), n detona la pregunta central: ¿cómo la implementación de Grok como herramienta de verificación de hechos (fact-checking) genera riesgos en la lucha contra la desinformación?

Este trabajo pretende analizar críticamente la función de Grok, la inteligencia artificial generativa integrada en la red sociodigital X, como herramienta de verificación de hechos,



para reflexionar sobre los riesgos, límites y dilemas éticos que implica delegar la construcción de la verdad a sistemas algorítmicos en el debate público digital.

Fact-checking e inteligencia artificial: el problema de los no-seres algorítmicos

El ecosistema informacional presenta sobreabundancia informativa. Por experiencia personal, quizás hayamos experimentado en algún momento una sensación de saturación informativa; más aún, de saberlo todo y nada al mismo tiempo. Porque en realidad así es.

Y es entre esta fauna informativa diaria con la que, siguiendo con la analogía, podemos toparnos con especies nocivas. Esta fauna nociva tiene distintos nombres, pero algunos de las más identificables son las fake news y la posverdad. Ambos conceptos aparecen en relevancia desde los sucesos de aquel 2020 cuando una pandemia puso a prueba nuestra capacidad de lidiar con un océano completo de información, pero también de desinformación.

Es pertinente detenernos a reflexionar por qué a veces es difícil discernir entre la información manipulada y los verdaderos hechos. Según Aparici y García (2019: 16) es porque, para mentir, debes partir de la verdad: “si la mentira no es verosímil, no funciona, deja de ser posverdad y se convierte en una mentira descubierta de entrada”. A ello se suma el complejo fenómeno de los prosumidores que, consciente o inconscientemente, se convierten en elementos de dispersión informativa y producen un efecto de teléfono descompuesto una vez que la fuente original se ha interpretado en diferentes plataformas y por diferentes cuentas de usuarios.

Es en esta realidad de fragilidad en la cual la verificación de los hechos, una acción originalmente ligada al campo de la investigación periodística, cobra una relevancia crítica. ¿De qué manera se ha modificado esta práctica en los espacios digitales? Mantzarlis (en Gutiérrez y Vázquez, 2024) dice que, a diferencia de la verificación periodística, “esta nueva forma de verificación de datos consiste en comprobar la información ya publicada y las declaraciones de actores relevantes en la esfera pública” (p. 56). Además de las cuentas que existen para la verificación de datos, como es el caso de VerificadoMx, la acción del fact-checking corre como responsabilidad o acción del usuario consciente e, incluso, es una característica importante de la alfabetización mediática. Lo problemático de esto es que,



como lo planteamos párrafos atrás, estamos ante una avalancha de fuentes, información y plataformas que hacen de la verificación de datos una tarea ardua para el usuario común (Gutiérrez y Vázquez, 2024).

Ante este panorama, la entrada de la IAG como auxiliar en esta acción parece una solución y “la IA tendría capacidad para leer el caos informacional (infoxicación) que existe en Internet y confirmar las noticias más dudosas, advirtiéndolo a los usuarios (lectores) cuáles de ellas pertenecen a la polémica categoría de “fake news” (Flores, 2019: 198). A diferencia de los verificadores humanos, las IAG no tienen información contextual ya que no interactúan con el mundo tangible. Por ello, carece de razonamiento causal. Muchas veces sus juicios resultan de modelos estadísticos y patrones aprendidos, no de una comprensión intrínseca o una intencionalidad (Bender et. al., 2021). Es aquí en donde radica la esencia del no-ser de la verdad. Los seres algorítmicos no poseen una verdad ontológica propia; su “verdad” es una construcción algorítmica; una inferencia basada en su entrenamiento que muchas veces permanece a modo de “caja negra” y que oculta sus procesos internos (O’Neil, 2016).

Grok, como IAG conectada a X, ejemplifica esta inserción algorítmica en la verificación. Su funcionamiento se basa en el acceso a la información que circula en tiempo real en la red, así como, presumiblemente, en grandes corpus de datos de entrenamiento. Esto nos lleva a interesantes interrogantes en cuanto se convoca a Grok a formar parte de una discusión, ya sea para realizar una confirmación de un hecho (“¿Grok, es esto real?”) como para tomar parte dentro de un acalorado tema político (“¿Grok, ¿qué piensas al respecto?”). ¿De qué manera se priorizan y filtran las fuentes en un entorno tan volátil como lo es X? Y muy importante: ¿cuáles son los criterios de definición que tiene para apoyar la tesis de un individuo que busca que Grok lo sustente en una acusación contra un político o un usuario?.

Hablando de responsabilidad: ¿finalmente quién es el dueño de la verdad?

Para los usuarios medianamente experimentados de las inteligencias artificiales generativas nos son familiares términos como alucinación o sesgo de confirmación, que saltan a la vista cuando utilizamos algún chatbot como ChatGPT para realizar consultas del día a día. Si bien estas “malfunciones” propias de los sistemas de IAG en muchos casos nos parecen inofensivas, cuando hablamos de deslindar responsabilidades en un caso de relevancia para la opinión pública el panorama se torna un poco más serio.



En el periodismo tradicional, los garantes de la verdad son el periodista, el editor y la institución mediática, cuya reputación y credibilidad están intrínsecamente ligadas a sus verificaciones de datos. Sin embargo, esta clara atribución de responsabilidad, fundamental para la credibilidad mediática, se difumina con la irrupción de la IA:

Los dispositivos y sistemas dotados de IA interactúan cada vez más con las personas y su entorno. Esta interacción suscita la pregunta de quién es responsable de los daños producidos en caso de que alguno de estos dispositivos opere erróneamente o tome una decisión de forma autónoma que resulte en algún tipo de perjuicio (Marin, 2019: 15).

Como ejemplo, imaginemos que Grok emite un juicio sobre la veracidad de una información en X que resulta ser incorrecta o perjudicial. Las preguntas sobre quién asume la responsabilidad se tornarían intrincadas: ¿es responsable el equipo que diseñó y entrenó al algoritmo, quizás introduciendo sesgos no intencionales? ¿Recae la responsabilidad sobre la regulación de X como plataforma al dejar integrar libremente a Grok como un usuario más en el debate?

En el complejo ejercicio de la verificación algorítmica, la figura del dueño de la verdad se desvanece en incertidumbre. Si en el periodismo formal la cadena de responsabilidad es clara y se cimienta en la ética y la reputación humana, la IA fragmenta esta línea transformando los daños informativos, que antes eran atribuibles, en una historia de sesgos y prejuicios que difícilmente podremos adjudicar a alguien o algo. Ante las prácticas de verificación a través de Grok nos encontramos en un debate con múltiples capas; pero, la más importante de todas es que la noción de la verdad, al ser mediada por un “no-ser” algorítmico, se vuelve una construcción opaca y mutable, y deja a la sociedad ante el dilema de redefinir finalmente quién o qué puede ser finalmente el garante de la verdad en la era digital.

El desafío de los “no-seres” algorítmicos

Estamos ante un posible desplazamiento del ejercicio de la verificación de hechos como una actividad humana hacia las IAG, con Grok como símbolo de esta nueva fase de tránsito a un ecosistema mediático donde la lógica estadística de un sistema diseñado para calcular probabilidades logra tener preponderancia para comprender significados y aproximarse a asuntos tan complejos como la valoración de la verdad, la autoridad y la construcción de la información pública.

Planteamos que estos elementos de diálogo con “no-seres” algorítmicos suponen retos importantes para el pensamiento crítico, el ejercicio de contrastar fuentes de información y construir de forma argumentada posturas ante el mundo. Grok no posee una intencionalidad moral ni una experiencia del mundo, y podría decirse que, “su verdad”, en última instancia, es una simulación probabilística del consenso, no una construcción racional basada en juicio o responsabilidad.

La noción de la verdad, tan importante para el campo periodístico, tiene en este contexto mayores responsabilidades con su labor ética, de investigación y de construcción de la opinión pública, ya que delegar la veracidad de la información a estos mecanismos automáticos opera como una solución aparente ante la desinformación, pero al mismo tiempo, genera nuevas zonas de ambigüedad.

El análisis del fact-checking algorítmico instauro una paradoja elemental, y al mismo tiempo, el principal reto que compete a las entidades humanas: mientras más se busca la certeza, más se deposita la confianza en entidades que carecen de la capacidad misma de experimentarla.

Fuentes

- Aparici, R., & García, A. (2019). *Desinformación, fake news y posverdad*. Gedisa.
- Bender, E. M., Gebru, T., McMillan-Major, A., & Shmitchell, S. (2021, 1 de marzo). On the dangers of stochastic parrots: Can language models be too big? [Conferencia]. *ACM Conference on Fairness, Accountability, and Transparency*, Canadá. <https://doi.org/10.1145/3442188.3445922>
- Flores, A. (2019). Inteligencia artificial y periodismo: diluyendo el impacto de la desinformación y las noticias falsas a través de los bots. *Doxa Comunicación*, 29, 197-212.
- Gutiérrez, B., & Vázquez, J. (2024). Redrawing the lines against disinformation: How AI is shaping the present and future of fact-checking. *Tripodos*, 55, 55-74.

- Marín, S. (2019). *Ética e inteligencia artificial* (Cuadernos de la Cátedra CaixaBank de Responsabilidad Social Corporativa, N° 42). <https://dx.doi.org/10.15581/018.ST-522>
- Musk, E. (2023, 17 de abril). Elon Musk to develop 'TruthGPT' artificial intelligence, warns of 'civilizational destruction' if AI isn't regulated. *Fox News*. <https://www.foxnews.com/media/elon-musk-develop-truthgpt-warns-civilizational-destruction-ai>
- O'Neil, C. (2016). *Weapons of math destruction: How big data increases inequality and threatens democracy*. Crown.
- X Corp. (2025). Acerca de Grok, tu divertido asistente de inteligencia artificial en X. <https://help.x.com/es/using-x/about-grok>