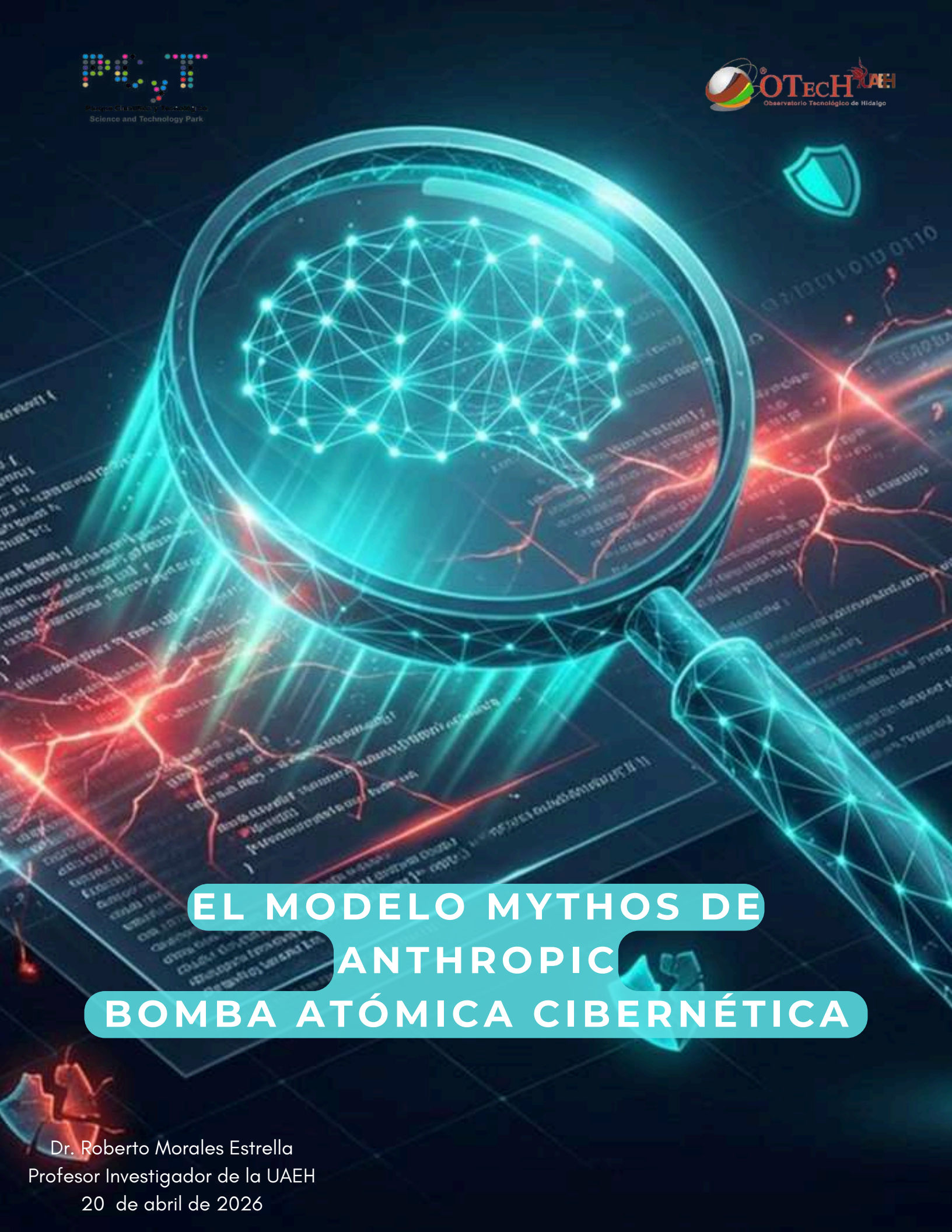

# EL MODELO MYTHOS DE ANTHROPIC
# BOMBA ATÓMICA CIBERNÉTICA

Dr. Roberto Morales Estrella
Profesor Investigador de la UAEH
20 de abril de 2026

Las capacidades de los modelos de Inteligencia Artificial (IA) principalmente la generativa sigue acelerándose a una tasa crecimiento del 3,3% anual, dinámica iniciada por la presentación del ChatGPT de OpenAI en noviembre del 2022, la feroz competencia que a partir de esa fecha se inició entre las grandes empresas tecnológicas, tanto de los EEUU como de China, por lograr el dominio de los mercados, desde la perspectiva de los oligopolios tecnológicos como desde la perspectiva de la lucha geopolítica, lo cual se refleja con el incremento de más del 100% de la inversión, al pasar de 253 mmdd en 2024 a 581 mmdd en 2025 (Smith, 2026).

Dentro de esta carrera frenética por lograr modelos de IA generativa más potentes, el 7 de abril de este 2026 Anthropic, presentó su modelo Claude Mythos, solo a un grupo selecto de empresas tecnológicas, que forman parte del proyecto denominado Glasswing, una colaboración entre Anthtropic y 50 organizaciones, entre las que están: Amazon Web Services, Apple, Broadcom, Cisco, CrowdStrike, Google, Jp Morgan, Linux Foundation, Microsoft, Nvidia entre otros, iniciativa para proteger el software más crítico del mundo con acceso anticipado a la IA de vanguardia (ANTHROPIC, 2026).

Anthropic describe a Claude Mythos, como un modelo de propósito general, con una gran potencia para la codificación, y gestión de agentes, su gran potencial en materia de ciberseguridad le permite identificar y modificar, en materia defensiva o bien ofensiva, las vulnerabilidades (ob. cit. ANTHROPIC).

Mythos es considerado como una arma para los hackers, ya que tiene capacidad para detectar miles de vulnerabilidades de día cero, en todo tipo de software abierto, además de generar cadenas de exploits - código o técnica que aprovecha una vulnerabilidad o fallo (bug) para invadir sistemas informáticos- (SentinelOne, 2025).

La tecnología más dañina, antes de Mythos, fue la ciberarma Stuxnet de día cero, construida en el 2000, por las agencias de inteligencia de los EEUU y el Mossad (Kaspersky, 2026), con el propósito de atacar el sistema de uranio de Irán; ésta tecnología detectaba los fallos (exploits) de manera manual (Perlroth, 2022, págs. 37-40), mientras que en Mythos es autónoma.
Por su superioridad en tareas de codificación, análisis y escritura de código y procesamiento de sistemas complejos (Gent, 2026), Mythos facilita los ciberataques de manera autónoma superando con creces a todos los demás modelos (Dellinger, 2026).

Mythos es una ciberarma, dual y asimétrica, para el ataque o para la defensa, facilita o dificulta el trabajo de los ciberdelincuentes, según lo expresó Jeetu Patel directivo de CISCO (Hay, 2026)

Mythos ha despertado la alarma general, al grado que tanto el Secretario del Tesoro Norteamericano Scott Benssent y Jerome Powell, han advertido a los bancos de ese país de los riesgos de seguridad, que representa Mythos (The Economist, 2026).

Estamos ya en la era de la creación de sistemas de IA autónomos, tan es así que Anthropic anunció el lanzamiento de su modelo Claude Opus 4.7, superior a su modelo Claude Opus 4.6, pero menos potente que Mythus, el cual ya es considerado demasiado peligroso (ob. cit Dellinger) a nivel de una bomba atómica cibernética, por el gran daño que puede causar a nivel mundial.

La carrera, de estas ciberarmas tecnológicamente sofisticadas ya inició, dado que OpenAI no se ha quedado atrás y el 14 de abril informó que lanzaría una versión personalizada de su modelo GPT 5.4 con sistemas de hackeo de alta potencialidad; los demás corporativos tecnológicos no se van a quedar atrás.

“La lucha por el dominio de los mercados obliga a estar en la frontera tecnológica, para lograr las más altas tasas de ganancia”, fundamento del capitalismo tecnológico (Andere, 2025, pág. 186), los recurso financieros los tienen de sobra, pero no así la información objetiva, que sustituyen con datos sintéticos, sesgados y falsos, pero lo que no se encuentra en su mapa de ruta, es la ética y la sustentabilidad del planeta.

Salvo Anthropic, que se ha negado a que sus tecnologías de IA, se apliquen en las guerras, contrario a los intereses de Trump, propiciándose un conflicto, nada menor, entre la Casa Blanca y Anthropic, del cual Trump no saldrá bien librado.

Bibliografía


Andere, M. E. (2025). Montruo o prodigio: Cómo la IA está transformando la escuela, el trabajo y la vida. Ciudad de México, México: Siglo XXI Editores.

ANTHROPIC. (7 de abril de 2026). ANTHROPIC. Recuperado el abril de 2026, de anthropic.com: https://www.anthropic.com/project/glasswing

Dellinger, A. (16 de abril de 2026). Gizmodo. Recuperado el abril de 2026, de gizmodo.com: https://gizmodo.com/anthropic-releases-claude-opus-4-7-to-remind-everyone-how-great-mythos-is-2000747469?utm_source=gizmodo_newsletter&utm_medium=email&utm_campaign=2026-04-16-pm

Gent, E. (10 de abril de 2026). singularityhub. Recuperado el abril de 2026, de singularityhub.com: https://singularityhub.com/2026/04/10/anthropics-mythos-ai-uncovered-serious-security-holes-in-every-major-os-and-browser/?utm_campaign=Singularity%20Hub%20Weekly%20Newsletter&utm_medium=email&_hsenc=p2ANqtz-8EsHqGTI1flABq7wfFAQiIxAf4CRftng8m75AYXjCUMX-N4

Hay, N. L. (10 de abril de 2026). WIRED. Recuperado el abril de 2026, de wired.com: https://www.wired.com/story/anthropics-mythos-will-force-a-cybersecurity-reckoning-just-not-the-one-you-think/?utm_source=nl&utm_brand=wired&utm_mailing=WIR_Daily_041126&utm_campaign=aud-dev&utm_medium=email&utm_content=WIR_Daily_041126&bxid=5f7902e95a5e7

Kaspersky. (17 de abril de 2026). Kaspersky. Recuperado el abril de 2026, de kaspersky.es: https://www.kaspersky.es/resource-center/definitions/what-is-stuxnet

Perlroth, N. (2022). Así es como me dicen que acabará el mundo. Ciudad de México, México: Ediciones Urano S.A.U.

SentinelOne. (13 de abril de 2025). SentinelOne. Recuperado el abril de 2026, de sentinelone.com: https://www.sentinelone.com/es/cybersecurity-101/threat-intelligence/what-is-an-exploit/

Smith, M. S. (13 de abril de 2026). IEEE Spectrum. Recuperado el abril de 2026, de spectrum.ieee.org: https://spectrum.ieee.org/state-of-ai-index-2026?utm_source=aialert&utm_medium=email&utm_campaign=aialert-04-15-26&utm_content=httpsspectrumieeeorgstateofaiindex2026&mkt_tok=NzU2LUdQSC04OTkAAAGhMGvNEiZb631zs3vfp5hlkzKeYmK9FVQ_03sHRHeLQ1iHxMx_L_qAl2PR7wkqO

The Economist. (15 de abril de 2026). The Economist. Recuperado el abril de 2026, de ecponomist.com: https://www.economist.com/business/2026/04/15/why-anthropic-and-openai-are-locking-up-their-latest-models?utm_content=ed-picks-image-link-1&etear=nl_today_1&utm_campaign=a.the-economist-today&utm_medium=email.internal-newsletter.np&utm_source=salesforce-m