FARO
INTELIGENCIA PARA LA EXCELENCIA OPERACIONAL

MARIO DEHTER
mariodehter.com

EDICIÓN ESPECIAL · 9 DE NOVIEMBRE DE 2026

La IA gana capacidad; la verificación, aún no

¿Quién comprueba lo que dicen los modelos?

El desplome del Nasdaq, una nueva unidad de Justicia y un test de agentes apuntan a lo mismo.

01
MERCADOS
El Nasdaq 100 cae un 6,4 % en un día

02
EE. UU.
Justicia crea la Super Intelligence Task Force

03
AGENTES
El mejor agente de código completa solo el 26,4 %

TOP 3 TRENDING HOY

REGULACIÓN
La UE reunirá a sus 27 líderes en Dublín solo para hablar de IA

Brussels Times · carta de Costa del 3 nov; cita el 13 nov

EDUCACIÓN · COLEGIOS (K-12)
Google: los colegios pueden apagar Gemini y conservar Search, que sigue con IA

FingerLakes1 · 8 nov

OPERACIONES
NVIDIA propone agentes de fábrica; cada planta fija qué decisiones exigen aprobación humana

MarketScale · 8 nov, según NVIDIA

LA CLAVE DEL DÍA

Las cifras clave las publica quien vende el modelo: la verificación independiente ya no es opcional.

MARIO DEHTER
mariodehter.com

MERCADOS · INFRAESTRUCTURA

El Nasdaq 100 se hunde un 6,4 % en un día por miedo a frenos a la IA

El jueves 5 de noviembre el Nasdaq 100 cayó un 6,4 %, su peor sesión desde 2020 según Swissinfo, y el índice de semiconductores perdió un 12 %. La prensa financiera lo atribuye al temor a que la victoria demócrata en las legislativas frene los centros de datos.

¿POR QUÉ IMPORTA?

El mercado trató la política de centros de datos como un riesgo para la inversión en IA, más allá de los modelos.

CON CAUTELA

Las cifras difieren: 2,3 billones de dólares (Swissinfo) o 2,5 (Devdiscourse). Una gestora ve un reajuste, no el fin de la inversión (9 nov).

FRASE CLAVE

Los centros de datos ya son un asunto político, además de técnico.

Fuentes: Swissinfo · Devdiscourse · Rallies.ai (Europe Capital Partners, 9 nov)

MARIO DEHTER
mariodehter.com

REGULACIÓN · EE. UU.

Justicia de EE. UU. crea la Super Intelligence Task Force

El 5 de noviembre el Departamento de Justicia anunció una unidad que coordinará la acción penal y civil contra el uso indebido de IA por Estados extranjeros, bandas y estafadores. Absorbe el grupo creado en enero para recurrir leyes estatales de IA y trabajará con la Super Intelligence Force de la Casa Blanca, según el fiscal general.

¿POR QUÉ IMPORTA?

La aplicación de la ley sobre IA pasa a ser prioridad de Justicia. Los fiscales podrían tener que avisar a la unidad antes de presentar cargos.

CON CAUTELA

The Neuron señala que sus dos tareas pueden chocar: vigilar abusos y frenar normas estatales.

FRASE CLAVE

Vigilar la IA y frenar las leyes estatales: dos misiones en una sola unidad.

Fuentes: GovExec · AI Weekly · The Neuron · Mind & Machine Weekly (8 nov)

MARIO DEHTER
mariodehter.com

AGENTES · FIABILIDAD

El mejor agente de código resuelve del todo solo el 26,4 % de tareas reales

Alibaba y varias universidades asiáticas publicaron Co-Design Bench, un test con peticiones reales de usuarios. GPT-6 Astra, el mejor de diez modelos, cumple de media el 70 % de los requisitos, pero completa del todo solo el 26,4 % de las tareas. En tests antiguos, los mismos modelos superan el 99 %.

¿POR QUÉ IMPORTA?

Los tests clásicos están saturados y ocultan fallos. Quien use agentes de código casi siempre debe repasar el resultado a mano.

CON CAUTELA

Cada tarea tiene una única petición fija, sin correcciones posteriores, y los jueces son IA que también pueden equivocarse.

FRASE CLAVE

Antes de elegir un agente, pruébalo con tus propias tareas y en tu propio entorno.

Fuentes: The Decoder (8 nov)

MARIO DEHTER
mariodehter.com

CIBERSEGURIDAD

OpenAI limita GPT-6 Cyber a usuarios verificados y lo califica «Crítico»

OpenAI lanzó GPT-6 Cyber el 5 de noviembre solo para clientes verificados de Daybreak Red. Según la empresa, su equipo logró con él 948 exploits de escalada de privilegios en Linux y halló 81 vulnerabilidades graves nuevas en Firefox. El 6 de noviembre, Singapur publicó un aviso conjunto (CSA, GovTech e IMDA) sobre el uso seguro de agentes de IA personales.

DATO CLAVE

Rating «Crítico» en el marco de preparación de OpenAI, un nivel por encima del «Alto» de GPT-5.6 Cyber.

CON CAUTELA

Las cifras son de la propia empresa; en las fuentes consultadas no consta verificación independiente.

FRASE CLAVE

Un modelo que encuentra fallos también puede explotarlos: por eso su acceso es limitado.

Fuentes: OpenAI · SecurityWeek · CSA Singapur (6 nov)

MARIO DEHTER
mariodehter.com

EDITORIAL

La pregunta ya no es qué pueden hacer los agentes, sino quién lo comprueba

Esta semana convivieron dos mensajes. Los modelos ganan capacidad: OpenAI presenta GPT-6 Cyber con 948 exploits para Linux, y METR estima, con mucha incertidumbre, que el ritmo de mejora de Anthropic casi se duplicó desde octubre de 2025. Pero el test con tareas reales dice otra cosa: el mejor agente de código completa el 26,4 %.

Buena parte de las cifras clave sale de quien vende el modelo, o de auditorías con la evidencia más informativa tachada: según METR, Anthropic tachó la mayor parte. Mientras, el mercado se asusta, Justicia crea una unidad y la UE reúne a sus líderes en Dublín el 13 de noviembre. Medir bien pasa a pesar tanto como construir.

LA GRAN PREGUNTA

¿Quién verifica al verificador?

Si los datos los publica quien vende el modelo, la verificación independiente deja de ser opcional. Y METR admite que depender de su criterio no es lo ideal.

EN RESUMEN

- **El Nasdaq 100 cayó un 6,4 % el 5 nov por el temor a frenos a los centros de datos.**
- **Justicia de EE. UU. creó la Super Intelligence Task Force y absorbe el litigio contra leyes estatales.**
- **El mejor agente de código completa del todo solo el 26,4 % de tareas reales.**
- **OpenAI limita GPT-6 Cyber a usuarios verificados y lo califica «Crítico».**

FUENTES: swissinfo.ch y devdiscourse.com (5 nov) · rallies.ai (9 nov) · govexec.com, aiweekly.co y theneuron.ai (5-6 nov) · mindandmachineweekly.substack.com (8 nov) · the-decoder.com (6 y 8 nov) · openai.com/index/introducing-gpt-6-cyber (5 nov) · securityweek.com · csa.gov.sg (6 nov) · brusselstimes.com (4 nov) · fingerlakes1.com (8 nov) · marketscale.com (8 nov)