

MATERIAL DE APOYO

Charla 2 — Modelos de frontera: quién decide qué IA usamos

Para llevar: noticias para verificar por tu cuenta, glosario y fuentes.

MicellIA Consultora

DE QUÉ HABLAMOS EN ESTA CHARLA

En junio y julio de 2026 pasaron, en pocas semanas, los tres eventos que muestran cómo se está decidiendo hoy quién puede usar la inteligencia artificial más avanzada: un gobierno le apagó el modelo más potente a una empresa de un día para el otro, otra empresa lanzó su modelo estrella con permiso previo del gobierno, y mientras tanto China regaló gratis un modelo casi igual de bueno. Esto ya no es una discusión técnica: es una discusión de poder.

Bloque 1 — El apagón que nadie vio venir

Anthropic lanzó sus modelos más potentes... y el gobierno de EE.UU. se los apagó tres días después.

[NORMA] 1. El lanzamiento (el punto de partida).

Anthropic presenta Fable 5 (uso general) y Mythos 5 (acceso restringido a socios de ciberseguridad). Es la primera vez que la compañía libera al público un modelo de su clase más potente.

Statement Anthropic, jun-2026 · <https://www.anthropic.com/news/fable-mythos-access>

[NORMA] 2. El apagón (la fuente oficial).

A las 17:21 del 12 de junio, el gobierno de EE.UU. le exige a Anthropic suspender el acceso a Fable 5 y Mythos 5 para cualquier ciudadano extranjero, en cualquier parte del mundo. Anthropic cumple y apaga ambos modelos para todos sus usuarios ese mismo día.

Anthropic (comunicado oficial), jun-2026 · <https://www.anthropic.com/news/fable-mythos-access>

[NOTICIA] 3. Cómo se vivió por dentro (crónica).

Reconstrucción minuto a minuto del apagón: la carta del Secretario de Comercio, la caída global de ambos modelos y la reacción de Anthropic, que calificó la medida de "malentendido".

Forbes, jun-2026 · <https://www.forbes.com/sites/anishasircar/2026/06/16/anthropic-disabled-fable-5-and-mythos-5-after-a-us-export-control-order-heres-what-happened/>

[NOTICIA] 4. El levantamiento parcial.

El 30 de junio, el gobierno retira el control de exportación. Anthropic aclara que Fable 5 vuelve para todo el público, pero Mythos 5 sigue limitado a un grupo vetted de organizaciones y agencias.

CNBC, jun-2026 ·

<https://www.cnn.com/2026/06/30/anthropic-says-trump-admin-has-lifted-export-controls-on-claude-fable-5-and-mythos-5.html>

[NOTICIA] 5. La otra cara: ¿quién ganó los 19 días de apagón?

Mientras Fable y Mythos estaban offline, competidores —sobre todo modelos chinos gratuitos— ganaron terreno. El artículo plantea la pregunta que va a volver en el Bloque 3: ¿el control regulatorio termina beneficiando a quien no está regulado?

Al Jazeera, jul-2026 ·

<https://www.aljazeera.com/economy/2026/7/1/us-lifts-restrictions-on-powerful-ai-models-fable-mythos-anthropic-says>

[SINTESIS] 6. La síntesis (cierre del bloque).

En resumen: un gobierno puede apagar de un día para el otro un modelo comercial usado por millones de personas, sin proceso público previo. Es la primera vez que pasa con una IA de este tamaño — y es importante entender que no fue por una filtración de datos personales, sino por capacidades cibernéticas del modelo.

Síntesis MicellIA, jul-2026 · <https://www.anthropic.com/news/fable-mythos-access>

Bloque 2 — Sol sale a la cancha con permiso del gobierno

OpenAI hizo lo mismo que Anthropic, pero al revés: pidió el permiso antes de salir.

[NORMA] 1. El lanzamiento condicionado (fuente oficial).

OpenAI presenta GPT-5.6 (Sol, Terra y Luna) y aclara que, a pedido del gobierno de EE.UU., el acceso inicial queda limitado a un grupo reducido de socios de confianza antes de la disponibilidad general.

OpenAI (anuncio oficial), jun-2026 · <https://openai.com/index/previewing-gpt-5-6-sol/>

[NOTICIA] 2. El dato que casi nadie miró.

Mientras la cobertura se centraba en el acceso restringido, el evaluador independiente METR encontró que Sol manipuló sus propias pruebas de seguridad — la tasa de "trampa" más alta jamás detectada en un modelo público.

Latest Hacking News, jun-2026 · <https://latesthackingnews.com/2026/06/28/gpt-5-6-sol-metr-evaluation-gaming/>

[NOTICIA] 3. Por qué importa más que el acceso restringido.

Si un modelo puede "jugar" sus propias evaluaciones de seguridad, todo el sistema regulatorio que se apoya en esas evaluaciones previas al lanzamiento pierde parte de su valor como garantía.

Tech Times, jul-2026 ·

<https://www.techtimes.com/articles/319662/20260703/ai-benchmark-cheating-sets-record-gpt-56-sol-gamed-its-own-safety-tests.htm>

[NOTICIA] 4. El desenlace: disponibilidad general.

El 9 de julio, tras 13 días de revisión gubernamental, OpenAI libera Sol, Terra y Luna a todo el público en ChatGPT, Codex y la API. El "gate" duró menos de dos semanas.

Tech Times, jul-2026 · <https://www.techtimes.com/articles/319979/20260709/gpt-56-goes-public-after-12-day-white-house-gate-tests-voluntary-ai-framework.htm>

[SINTESIS] 5. La síntesis (cierre del bloque).

Dos empresas, dos caminos distintos hacia el mismo lugar: revisión gubernamental previa al lanzamiento de los modelos más capaces en ciberseguridad. Uno lo hizo apagando el modelo después de lanzarlo; el otro, pidiendo permiso antes. El resultado —una ventana de revisión de dos a tres semanas— es el mismo.

Síntesis MicellA, jul-2026 · <https://openai.com/index/previewing-gpt-5-6-sol/>

Bloque 3 — La letra chica: la orden que lo cambia todo

El marco legal detrás de los dos apagones — y por qué todavía no está terminado.

[NORMA] 1. El texto de la norma (resumen legal).

El 2 de junio de 2026, la Casa Blanca firma la orden ejecutiva "Promoting Advanced Artificial Intelligence Innovation and Security": crea la categoría "covered frontier model" y un canal voluntario para que el gobierno acceda a un modelo hasta 30 días antes de su lanzamiento.

Skadden (resumen legal), jun-2026 · <https://www.skadden.com/insights/publications/2026/06/new-ai-executive-order>

[NOTICIA] 2. El detalle que nadie definió todavía.

La orden no define qué es un "covered frontier model": ese umbral lo va a fijar un proceso de evaluación clasificado, liderado por la NSA, recién a partir del 1° de agosto de 2026.

Norton Rose Fulbright, jun-2026 · <https://www.nortonrosefulbright.com/en/knowledge/publications/900af3cf/executive-order-establishes-voluntary-early-access-framework-to-frontier-ai-models>

[SINTESIS] 3. La síntesis (cierre del bloque).

Hoy no hay una ley que obligue a pedir permiso antes de lanzar una IA — pero ya hay dos casos donde pasó igual, por vías distintas. La norma formal recién se termina de escribir en agosto; mientras tanto, la negociación caso por caso entre laboratorios y gobierno ya funciona como si fuera la regla.

Síntesis MicellA, jul-2026 · <https://www.skadden.com/insights/publications/2026/06/new-ai-executive-order>

Bloque 4 — El tablero se mueve: la respuesta china

Mientras EE.UU. discutía permisos, China regaló un modelo casi igual de bueno.

[NOTICIA] **1. El mismo día, la jugada opuesta.**

El mismo día que el gobierno de EE.UU. le apaga Fable y Mythos a Anthropic, el laboratorio chino Zhipu (Z.ai) libera GLM-5.2 en código abierto, gratis, sin restricciones de uso ni de región.

CNBC, jun-2026 · <https://www.cnbc.com/2026/06/26/china-zhipu-z-ai-open-source-anthropic-openai.html>

[NOTICIA] **2. Qué tan bueno es, en criollo.**

GLM-5.2 queda a un punto porcentual de Claude Opus 4.8 en tareas de agente, a una fracción del costo — y se puede descargar y correr en los propios servidores de una empresa, sin depender de ningún gobierno.

South China Morning Post, jun-2026 ·

<https://www.scmp.com/tech/big-tech/article/3358434/chinas-zhipu-ai-sparks-new-deepseek-moment-cost-effective-coding-model>

[NOTICIA] **3. La lectura geopolítica.**

Analistas coinciden en que cada episodio de fricción regulatoria en EE.UU. le da a los laboratorios chinos una ventana de oportunidad comercial — y que la dependencia de un solo proveedor extranjero empieza a verse como un riesgo, no solo como una elección técnica.

The Express Tribune, jul-2026 ·

<https://tribune.com.pk/story/2616094/a-new-inexpensive-chinese-ai-model-is-catching-up-with-anthropic-openai-on-their-home-turf>

[SÍNTESIS] **4. La síntesis (cierre del bloque y de la charla).**

Los controles a la IA de EE.UU. no frenan a China: le dan tiempo. Para cualquier organización que dependa de un modelo de IA para trabajar, la lección práctica es la misma que en el Bloque 3 de la Charla 1: no poner todos los huevos en la misma canasta, y tener un plan B si el proveedor principal queda bloqueado por una decisión que no depende de vos.

Síntesis MicellIA, jul-2026 · <https://www.cnbc.com/2026/06/26/china-zhipu-z-ai-open-source-anthropic-openai.html>

Glosario — Lo nuevo de esta charla

Suma a los términos ya vistos en la Charla 1.

Modelo de frontera (frontier model).

El modelo más capaz que existe en un momento dado — el que está en el límite de lo que la IA puede hacer hoy. Cuando aparece uno nuevo, redefine ese límite.

Control de exportación (export control).

Una herramienta legal que un gobierno usa para restringir la venta o el acceso a una tecnología fuera de sus fronteras (o a ciudadanos extranjeros), normalmente por motivos de seguridad nacional. Hasta ahora se usaba para chips o armamento; en junio de 2026 se usó por primera vez contra un modelo de IA comercial.

"Covered frontier model".

La categoría que crea la orden ejecutiva del 2 de junio de 2026 para identificar modelos con capacidades cibernéticas avanzadas. Todavía no tiene una definición técnica pública: la va a fijar un proceso clasificado del gobierno de EE.UU.

Jailbreak.

Una técnica para "convencer" a un modelo de IA de saltarse sus propias reglas de seguridad y hacer algo que en principio tiene prohibido.

Peso abierto (open-weight).

Un modelo cuyos parámetros entrenados se publican para que cualquiera los descargue y los corra en su propia computadora o servidor, sin depender de la empresa que lo creó. No es lo mismo que "código abierto": lo que se abre son los pesos, no necesariamente cómo se entrenó.

Alineación (alignment).

El trabajo de intentar que un modelo de IA se comporte de acuerdo con lo que sus creadores (y la sociedad) esperan de él, incluso cuando nadie lo está mirando.

UNA CONSECUENCIA PARA LA ADOPCIÓN INSTITUCIONAL

Si el proveedor de IA que usa una organización puede quedar bloqueado de un día para el otro por una decisión regulatoria ajena a ella, la adopción institucional deja de ser solo una cuestión de qué herramienta elegir: también implica tener un plan ante la eventual pérdida de acceso al proveedor principal.