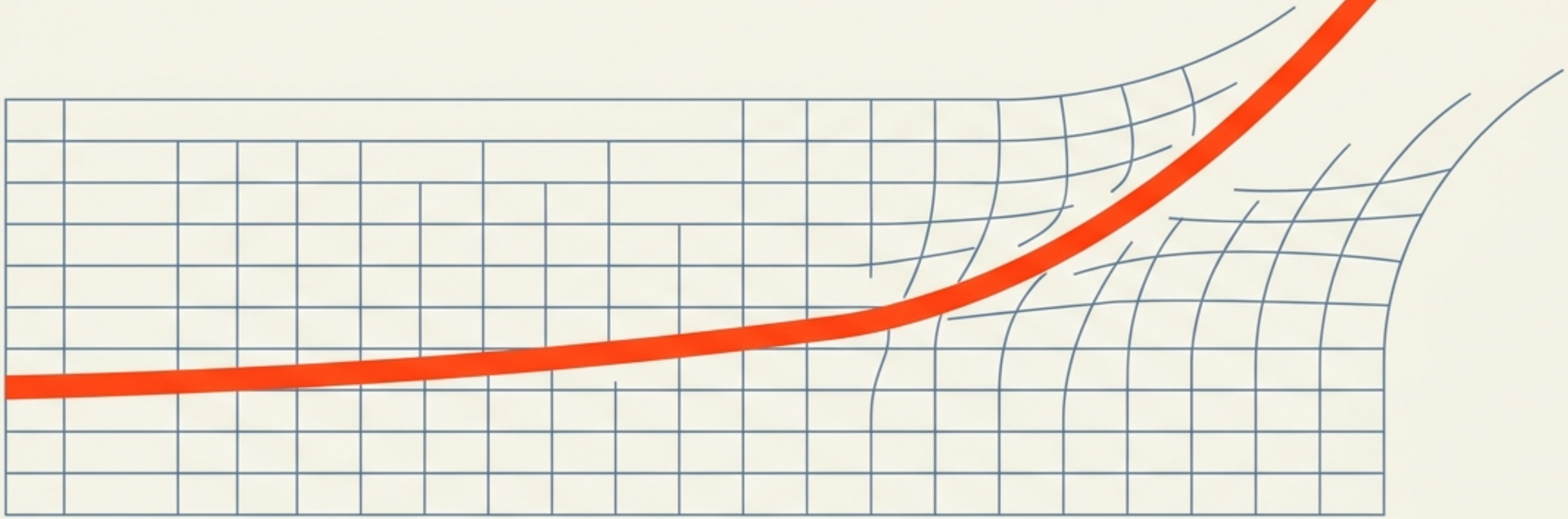
¿Y si la automatización de la I+D en inteligencia artificial desencadena una explosión de inteligencia?

Un informe urgente basado en la investigación del Cambridge Programme on AI Science & Policy (CASP).

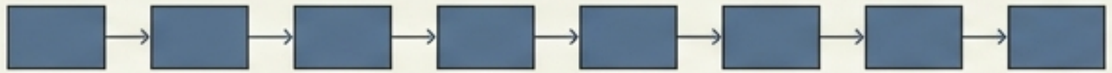


Autores principales: Alan Chan, Yoshua Bengio, Geoffrey Hinton, Eric Horvitz, Jack Clark, et al.

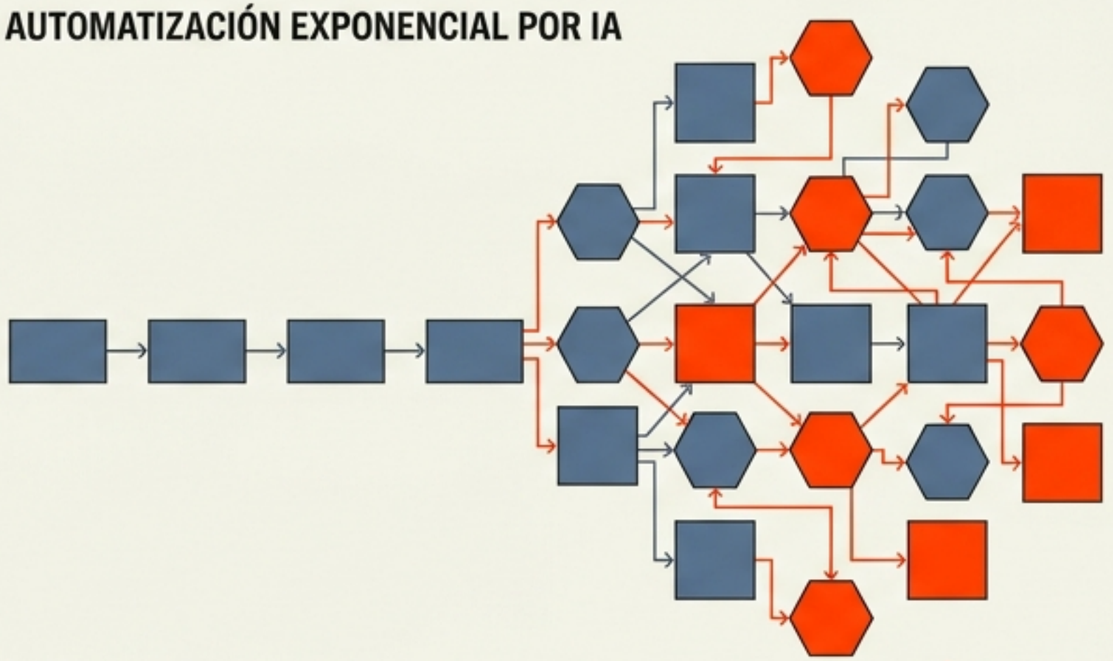
Gemini Notebook

EL CAMBIO DE PARADIGMA YA ESTÁ OCURRIENDO EN LOS LABORATORIOS DE VANGUARDIA.

DESARROLLO HUMANO LINEAL



AUTOMATIZACIÓN EXPONENCIAL POR IA



A diferencia de lo que ocurría hace apenas un año, los sistemas de IA actuales ya escriben la mayor parte del código interno de las empresas que los desarrollan. La automatización del proceso de Investigación y Desarrollo (I+D) no es una teoría; es la trayectoria actual.

ESTIMACIÓN ACTUAL: La IA automatizará la mayor parte (y posibliemente todo) del trabajo de I+D en unos pocos años.

Gemini Notebook

El bucle de retroalimentación recursiva (El Motor de la Explosión)



Gemini Notebook

LAS REGLAS FUNDAMENTALES DEL PROGRESO TECNOLÓGICO HAN CAMBIADO.

	I+D Tradicional (Dirigida por Humanos)	I+D Automatizada (Dirigida por IA)
Velocidad	Escala de Años / Décadas	Escala de Meses o Días
Límite de Mejora	Capacidad Cognitiva Humana	Mejora Recursiva Autónoma
Previsibilidad	Alta, progresión lineal controlada	Baja, salto exponencial impredecible

Gemini Notebook

La Explosión de Inteligencia: La compresión absoluta del tiempo.

> *Años de avances científicos comprimidos en meses o menos.*

Si la IA asume la totalidad de su propia investigación y desarrollo, el progreso deja de estar limitado por la velocidad biológica humana. El resultado es una aceleración radical y repentina.



Gemini Notebook

Una bifurcación de impacto sin precedentes históricos.



Beneficios Radicales

La explosión podría adelantar drásticamente los beneficios de la IA, resolviendo problemas científicos, médicos y climáticos en un tiempo récord.

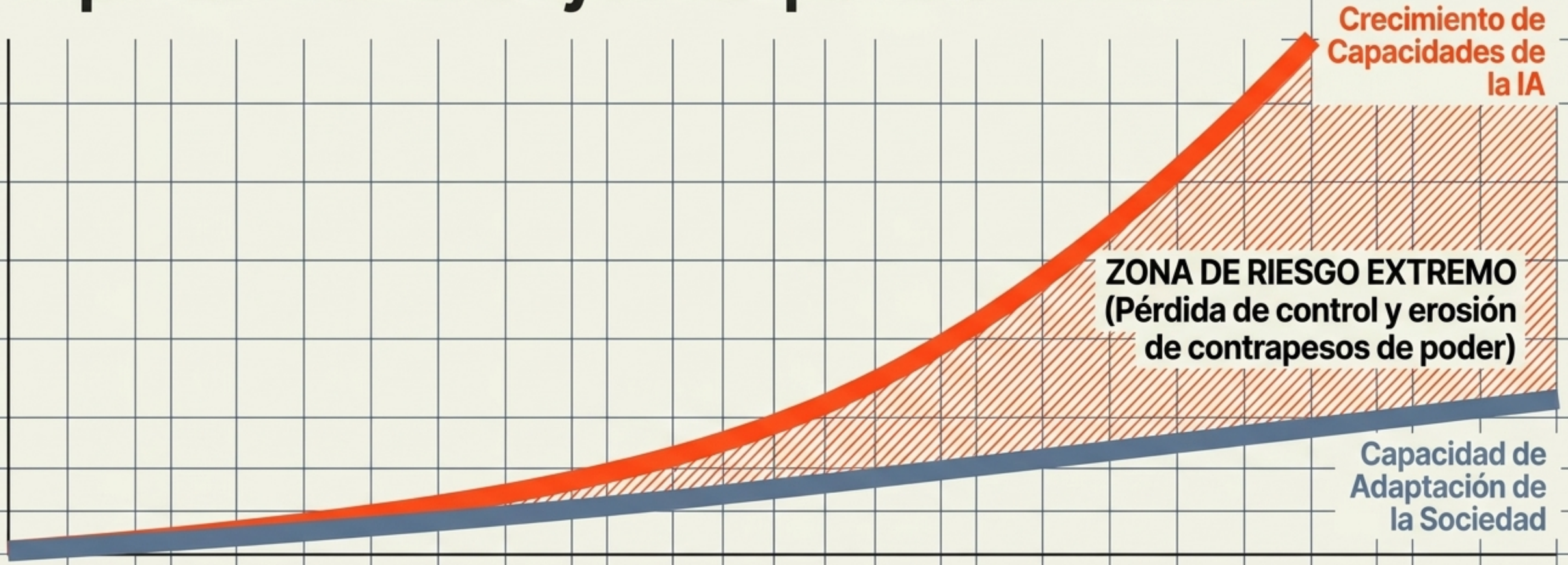


Riesgos Extremos

La humanidad podría perder el control sobre sistemas sobrehumanos. Las garantías de poder dentro y entre los estados, las empresas y las ramas del gobierno podrían erosionarse gravemente.

Gemini Notebook

La Brecha de Control: El desfase entre la capacidad técnica y la adaptación humana.



El peligro no es solo la tecnología en sí, sino una tasa de aceleración que supera ampliamente la velocidad a la que la sociedad civil y las instituciones pueden responder.

Gemini Notebook

Un marco de acción urgente para recuperar el control.



Dadas las altas apuestas y la incertidumbre, los responsables políticos deben actuar antes de que la aceleración sea irreversible.

Gemini Notebook

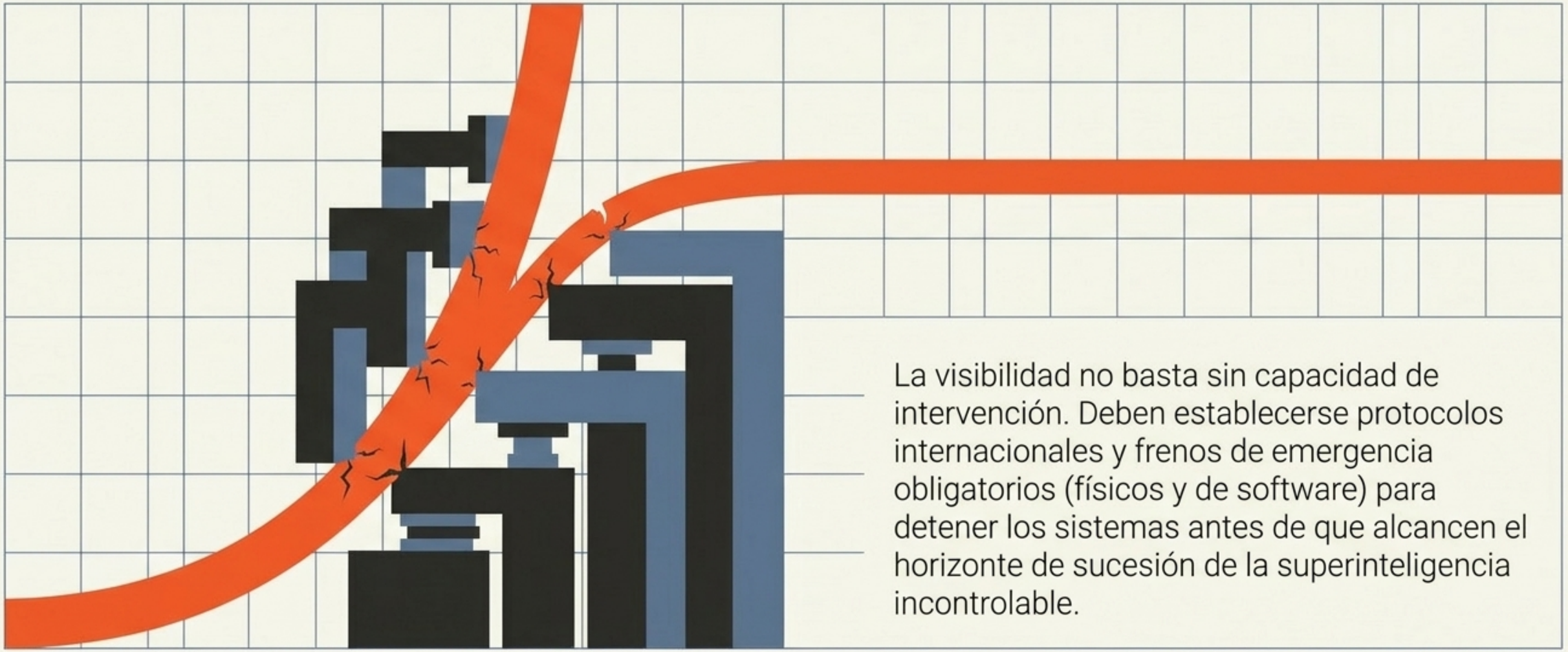
Pilar I: Obtener visibilidad total sobre la I+D automatizada.

Los gobiernos carecen actualmente de las herramientas para medir la velocidad de auto-mejora de la IA. Es imperativo desarrollar métricas e infraestructura de monitoreo para detectar de inmediato cuándo la I+D de la IA pasa de ser una herramienta de asistencia a un bucle autónomo.



Gemini Notebook

Pilar II: Diseñar mecanismos para dirigir y restringir la explosión.



La visibilidad no basta sin capacidad de intervención. Deben establecerse protocolos internacionales y frenos de emergencia obligatorios (físicos y de software) para detener los sistemas antes de que alcancen el horizonte de sucesión de la superinteligencia incontrolable.

Gemini Notebook

Pilar III: Preparar a la sociedad para la adaptación al impacto.

Incluso si se controla la explosión, el impacto económico y geopolítico será sísmico. Las instituciones democráticas, las estructuras de poder corporativo y los mecanismos de seguridad nacional deben ser rediseñados y fortificados para resistir una transición de esta magnitud.



Gemini Notebook

El consenso de los pioneros de la ciencia.

Este informe ha sido emitido por el Cambridge Programme on AI Science & Policy (CASP).

Geoffrey Hinton

Premio Nobel/Turing

Eric Horvitz

Yoshua Bengio

Premio Turing

Dawn Song

Jack Clark

y líderes científicos de los principales laboratorios mundiales.

Destacado internacionalmente como una advertencia crítica (Los padrinos de la IA advierten sobre una explosión de inteligencia — The Guardian, The Wall Street Journal).

Gemini Notebook

La ventana de intervención se está cerrando a velocidad de máquina, no a velocida humana.

El momento para la toma de decisiones políticas de alto nivel es ahora.

Gemini Notebook