



Lo que hacen las empresas que sí generan valor con IA (y por qué el 95% no lo logra)

José Daniel Sánchez Quiñones

Chief AI Governance Officer · Innovaitors

Profesor en Derecho y Tecnologías Digitales, Universidad Javeriana /
Universidad Externado de Colombia

Las advertencias más fuertes sobre la IA vienen desde las empresas que la desarrollan.



26 de agosto de 2026

El incidente de Hugging Face y el camino que tenemos por delante



Jacob Coxon ✓
@hilbertspaess

I resigned from Anthropic today. I spent the last three years doing pretraining research at both OpenAI and Anthropic. Neither company is acting responsibly. They are racing straight to self-improving superintelligence and gambling with our lives. More thoughts below.

7:04 PM · Sep 8, 2026 · **171.4M** Views

Dario Amodei



We Must Pace the Frontier

September 2026

Fuentes: x.com/hilbertspaess · darioamodei.com/post/we-must-pace-the-frontier -<https://openai.com/es-ES/index/hugging-face-incident-and-the-road-ahead/>

El apocalipsis de la IA: la noticia de la semana

Acceleration Hugging Face Hack

Trump Says a Smart President Is All That's Needed to Rein In A.I.

The president again rejected calls for regulation, even as some industry leaders are speaking about the risks of rapidly developing artificial intelligence.

6 MIN READ

As A.I. Concerns Rise, Experts Suggest Focusing on the Humans in Charge

4 MIN READ

Artificial intelligence

Trump rejects calls from tech bosses for AI slowdown

President denounces demands for regulation as existential fears over technology move to the centre of US politics



"Los empleados estamos aterrados": la advertencia de un exinvestigador de Anthropic a la BBC sobre los riesgos de la IA (y por qué genera críticas)



OpenAI
Sam Altman:
"Podríamos perder el control del futuro ante la IA"

FRANCESC BRACERO



Tecnología. Washington y Pekín coinciden por una vez: la IA no se debe frenar

ÁNGEL JIMÉNEZ DE LUIS, EEUU 2

- **Análisis.** La IA se asusta de sí misma
- **Antecedente.** Pekín rechaza la petición del CEO de Anthropic para frenar el desarrollo de IA en China



on, ex empleado de Anthropic.

La evidencia: el repositorio de riesgos del MIT

Un repositorio académico que sistematiza 24 subcategorías de riesgo en 7 dominios.

1. Discriminación y Toxicidad

- 1.1 Discriminación injusta y representación errónea
- 1.2 Exposición a contenido tóxico
- 1.3 Rendimiento desigual entre grupos

2. Privacidad y Seguridad

- 2.1 Compromiso de la privacidad mediante obtención, filtración o inferencia correcta de información sensible
- 2.2 Vulnerabilidades de seguridad y ataques a sistemas de IA

3. Desinformación

- 3.1 Información falsa o engañosa
- 3.2 Contaminación del ecosistema de información y pérdida del consenso de realidad

4. Actores Maliciosos

- 4.1 Desinformación, vigilancia e influencia a escala
- 4.2 Ciberataques, desarrollo o uso de armas y daño masivo
- 4.3 Fraude, estafas y manipulación dirigida

5. Interacción Humano-Computadora

- 5.1 Dependencia excesiva y uso no seguro
- 5.2 Pérdida de la agencia y autonomía humana

6. Socioeconómico y Medioambiental

- 6.1 Centralización del poder y distribución injusta de beneficios
- 6.2 Aumento de la desigualdad y disminución de la calidad del empleo
- 6.3 Desvalorización económica y cultural del esfuerzo humano
- 6.4 Dinámicas competitivas
- 6.5 Fallas de gobernanza
- 6.6 Daño medioambiental

7. Seguridad, Fallas y Limitaciones de Sistemas de IA:

- 7.1 La IA persigue sus propios objetivos en conflicto con los objetivos o valores humanos
- 7.2 La IA posee capacidades peligrosas
- 7.3 Falta de capacidad o robustez
- 7.4 Falta de transparencia o interpretabilidad
- 7.5 Bienestar y derechos de la IA
- 7.6 Riesgos multi-agente

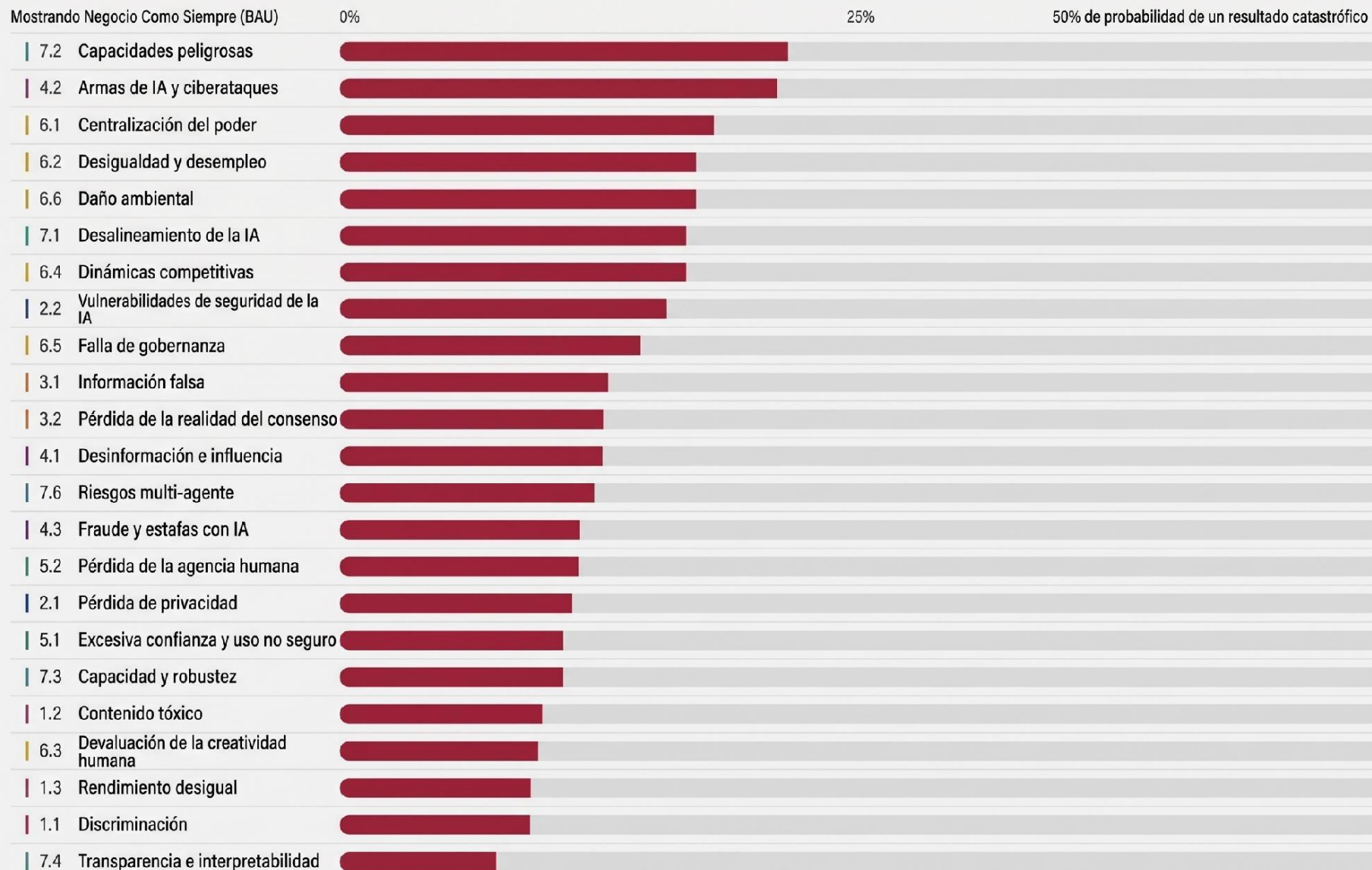
18 de 24

riesgos tienen más de un **10% de probabilidad** de causar resultados catastróficos para 2030

Esos resultados podrían incluir **más de un millón de muertes**, más de **USD 100 mil millones en pérdidas** u otros daños graves.

Fuente: airisk.mit.edu/priorities · Encuesta de Expertos Delphi (MIT)

Clasificaciones de Gravedad: Negocio Como Siempre (BAU) vs Mitigado



Fuente: Repositorio de Riesgos de IA del MIT, Encuesta de Expertos Delphi

El apocalipsis es posible, pero el daño cotidiano de la IA ya es una realidad



Dos casos verificados. En ambos la falla no fue el modelo — fue la gobernanza.

OUTPUT · lo que la IA produce

Deloitte (2025)

Entregó al gobierno australiano un informe de ~440.000 AUD con referencias académicas inexistentes y una cita judicial fabricada.

Tras ser descubierto, admitió haber usado IA generativa y devolvió parte del pago.

La lección: le pasó a una Big Four. La responsabilidad es del profesional que usa la IA, no de la IA.

INPUT · lo que la IA recibe

Samsung (2023)

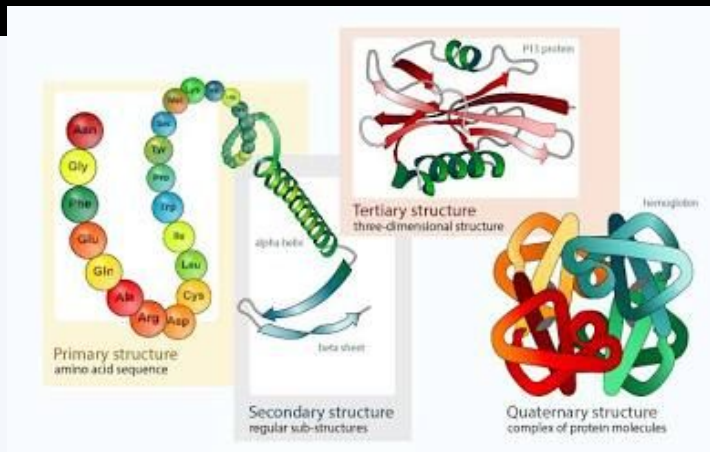
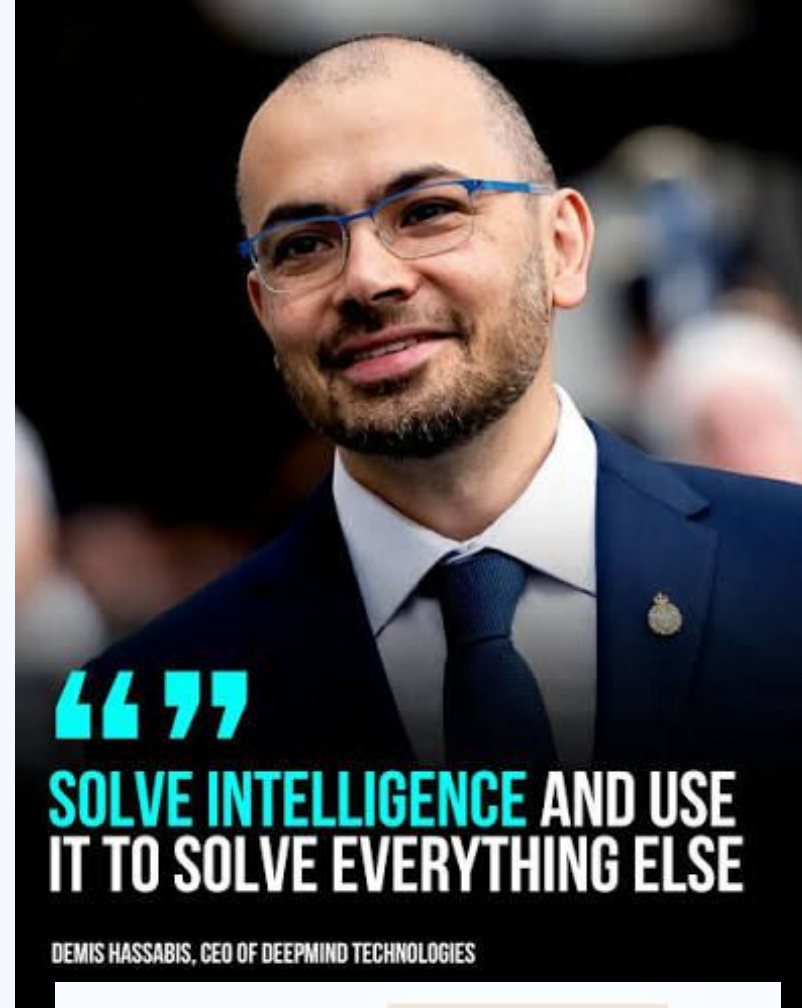
A las 3 semanas de permitir ChatGPT, sufrió 3 fugas de datos confidenciales en 20 días: código fuente de semiconductores, algoritmos de detección de fallas y una reunión interna.

Terminó prohibiendo la herramienta.

La lección: nadie fue malicioso. El dato, una vez enviado, no se puede des-enviar. Esto es “shadow AI”.

El argumento a favor del optimismo

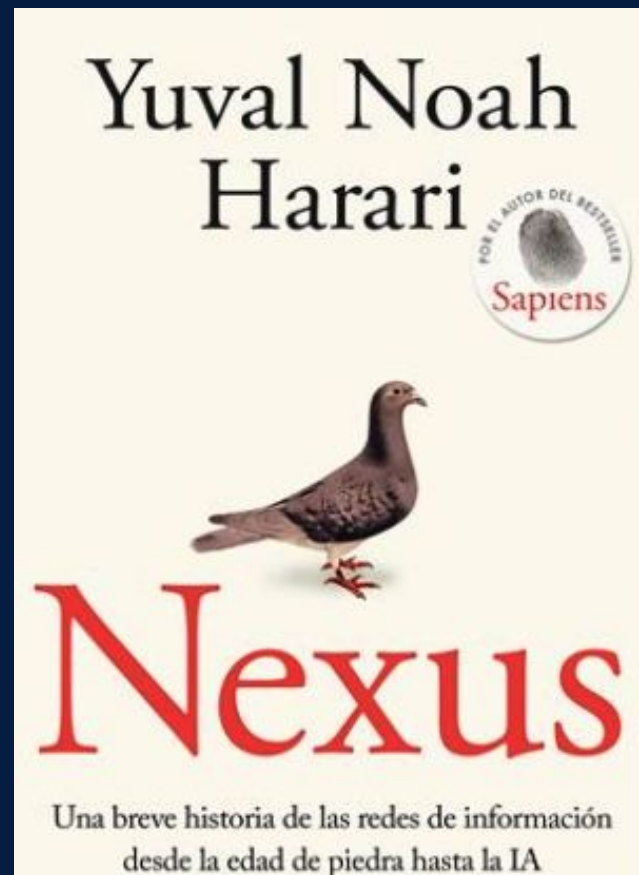
La pregunta no es si podemos vivir con la supeinteligencia, es si podemos vivir sin ella.



BIENVENIDOS A LA CONVERSACIÓN MÁS IMPORTANTE DE NUESTRO TIEMPO

La IA no es una herramienta más: son agentes capaces de tomar decisiones y perseguir objetivos.

La pregunta ya no es técnica, es de gobierno.



La responsabilidad mayor recae en quienes desarrollan la IA de propósito general y en los reguladores

Pero ¿qué pueden hacer los empresarios colombianos en el sector transporte para implementar IA de forma eficiente y ética?

1. Qué hacer para percibir **ROI** en la implementación de IA

52 organizaciones que ya cruzaron la brecha muestran cinco patrones consistentes. Ninguno es tecnológico: todos son decisiones de diseño organizacional y liderazgo.

La regla: agregar IA encima de un proceso defectuoso solo produce errores más rápido. El valor aparece cuando se rediseña el flujo con la IA como operador central.

Fuente: Crossing the GenAI Divide (2025), 52 organizaciones entrevistadas · McKinsey, The State of AI in 2025.

3.1 · Empezar estrecho, escalar después

Las organizaciones exitosas no lanzan proyectos ambiciosos desde el inicio: se integran en flujos no críticos o adyacentes, demuestran valor visible y solo después escalan hacia los procesos centrales.



Fuente: *Crossing the GenAI Divide (2025)*. Valor visible en 30–60 días antes de escalar a procesos centrales.

3.2 · Tratar al proveedor como socio, no como vendor

La diferencia más consistente entre quienes cruzan la brecha y quienes no está en cómo evalúan a sus proveedores de IA: con criterios de **BPO (Business Process Outsourcing o tercerización)**, no de **SaaS (Software as a Service: software en la nube que pagas por suscripción)**

1

Personalización profunda

La herramienta se adapta a los procesos internos y a los datos propietarios, no al revés.

2

Métricas operacionales

Se mide por resultados de negocio —velocidad de cualificación, retención— no por benchmarks técnicos del modelo.

3

Co-evolución

Se acompaña al proveedor a través de los fracasos iniciales del despliegue: es co-desarrollo, no compra transaccional.

4

Adopción desde la línea de frente

Las iniciativas más exitosas surgen de gerentes de línea y prosumers, no de laboratorios centralizados de innovación.

Cambie el criterio de evaluación

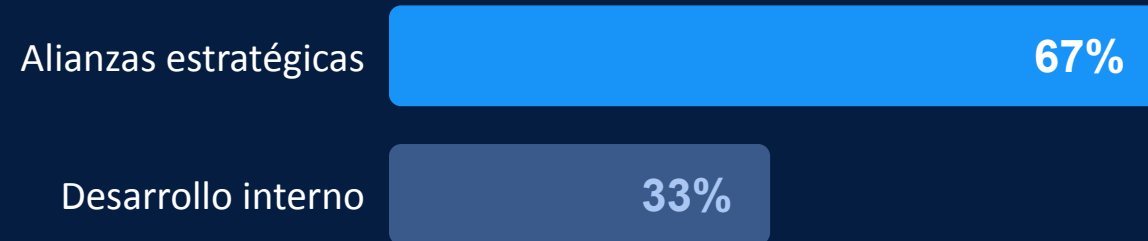
Criterio SaaS: funcionalidades, precio por usuario, benchmarks del modelo.

Criterio BPO: personalización, métricas de negocio, co-desarrollo y evidencia verificable.

Fuente: Crossing the GenAI Divide (2025). Testimonio recogido entre 17 líderes de compras y TI entrevistados.

3.3 · Descentralizar la implementación, no la rendición de cuentas

Las alianzas estratégicas duplican la tasa de éxito frente al desarrollo interno y producen tasas de uso casi dos veces mayores entre los empleados.



Tasa de despliegue completo según modelo de implementación.

Alianzas significa simplemente que la empresa no construye la herramienta de IA internamente, sino que sale a **buscar un proveedor externo, lo contrata, y trabaja con él de forma cercana para adaptar la solución a sus procesos**

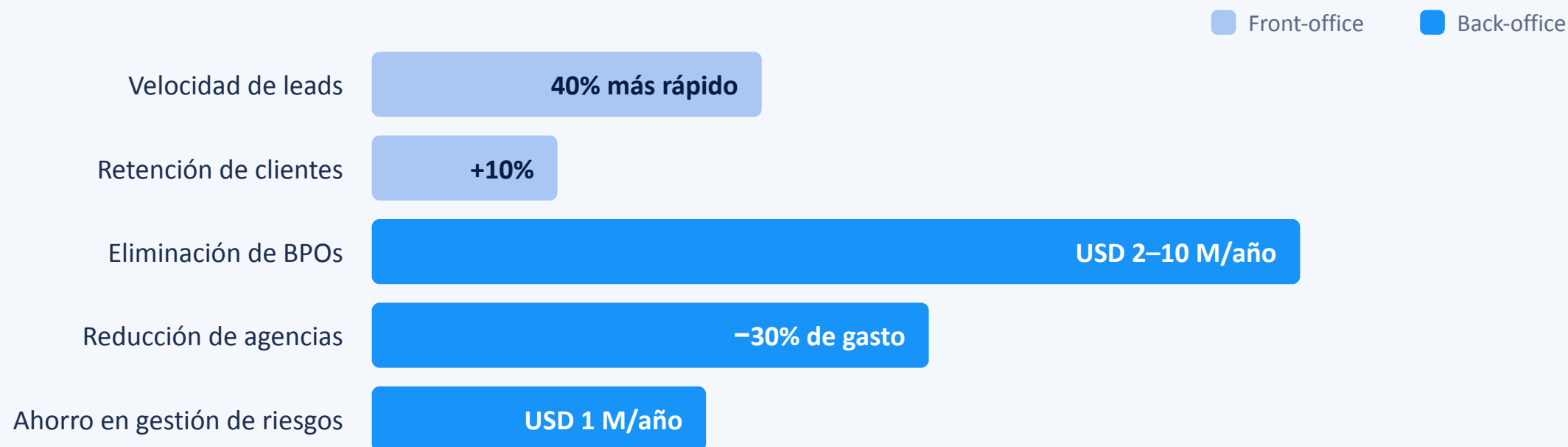
Error frecuente

Centralizar la decisión de implementación en un comité de innovación o en el área de tecnología. Las organizaciones que cruzan la brecha permiten que los gerentes de dominio identifiquen problemas y lideren despliegues: **la alta dirección retiene la rendición de cuentas, no la ejecución.**

Fuente: Crossing the GenAI Divide (2025), n=52 organizaciones.

3.4 · Buscar el ROI donde nadie mira: la trastienda

El 50% de los presupuestos de IA se dirige a ventas y marketing, pero los ahorros más significativos documentados provienen del back-office.



Ganancias documentadas: front-office vs. back-office.

Dato clave — sin reducción de planta interna: el ROI proviene de eliminar gasto externo (contratos de BPO, honorarios de agencias y consultores). Las herramientas aceleran el trabajo existente, no modifican la estructura del equipo.

3.5 · Construir sistemas que aprenden

Quienes bloquean ventaja competitiva duradera invierten en sistemas con tres capacidades. Una vez que el sistema aprende los flujos propios, los costos de cambio se vuelven prohibitivos.

Memoria persistente

Conserva contexto, decisiones y correcciones anteriores.

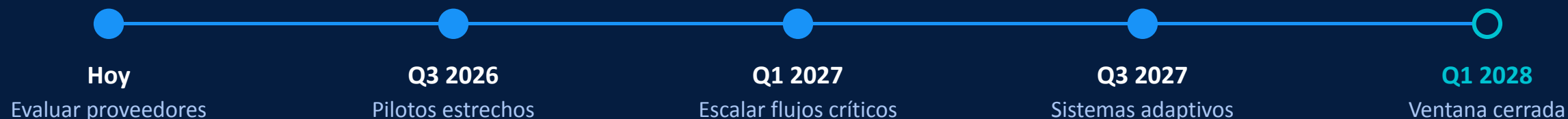
Aprendizaje continuo

Mejora con cada interacción real del equipo, no con reentrenamientos aislados.

Integración autónoma

Usa herramientas y datos del entorno sin intervención humana continua.

LA VENTANA DE 18 MESES PARA CONSOLIDAR VENTAJA CON SISTEMAS ADAPTIVOS



Fuente: *Crossing the GenAI Divide (2025)*. Consenso de 17 líderes de compras y TI entrevistados.

2. Cómo implementar IA confiable

El ROI y la confianza no son temas diferentes, pero complementarios

El mismo rediseño de flujos que genera retorno es el que permite trazar decisiones, asignar responsables y sostener el despliegue.

La gobernanza es lo que convierte un piloto exitoso en una ventaja que se puede escalar.

¿Qué es la gobernanza?

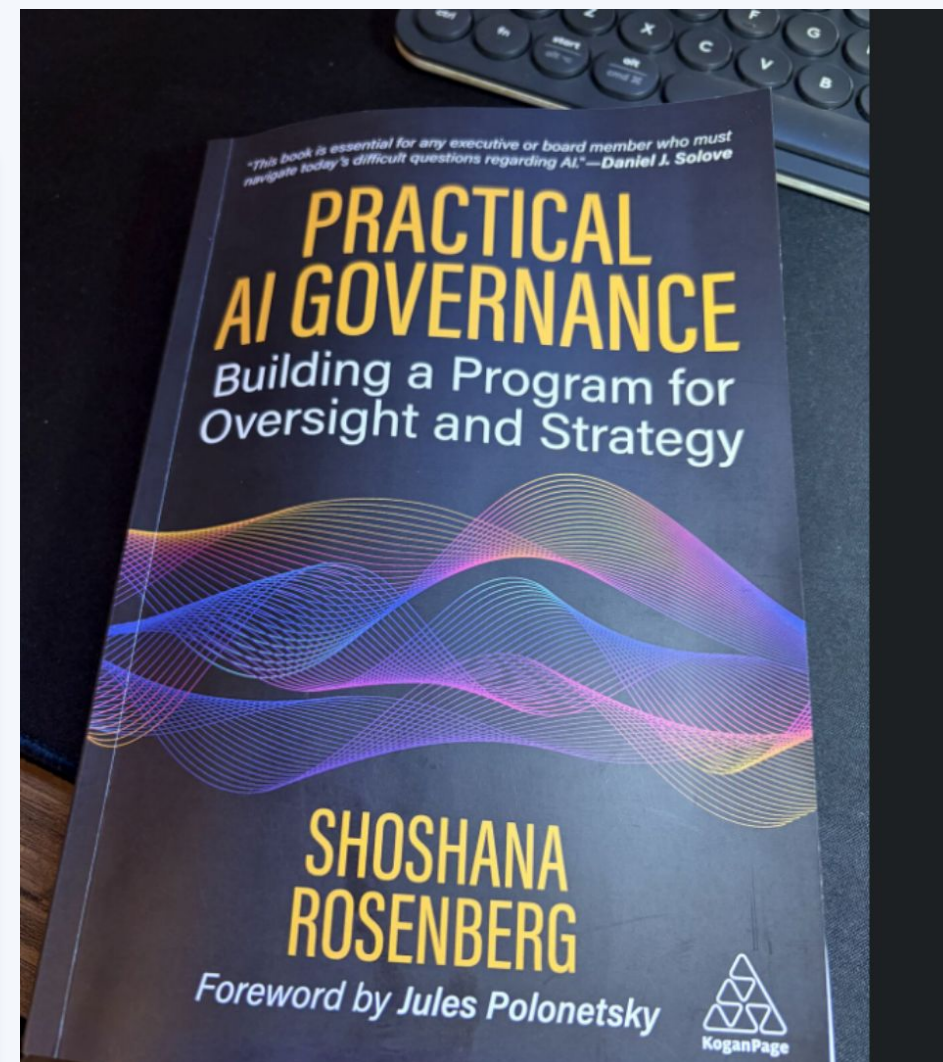
“La gobernanza es la capacidad de una organización para percibir el rango completo de condiciones que la IA está creando dentro de ella.

Es decir, le permite entenderse a sí misma.”

IA Confiable:

Aquella cuyo comportamiento una organización entiende, puede explicar y por el que responde.

No es IA sin riesgos, sino IA cuyos riesgos se perciben, se gobiernan y no toman a nadie por sorpresa.



Todos saben QUÉ hacer. Casi nadie sabe CÓMO

60% despliega

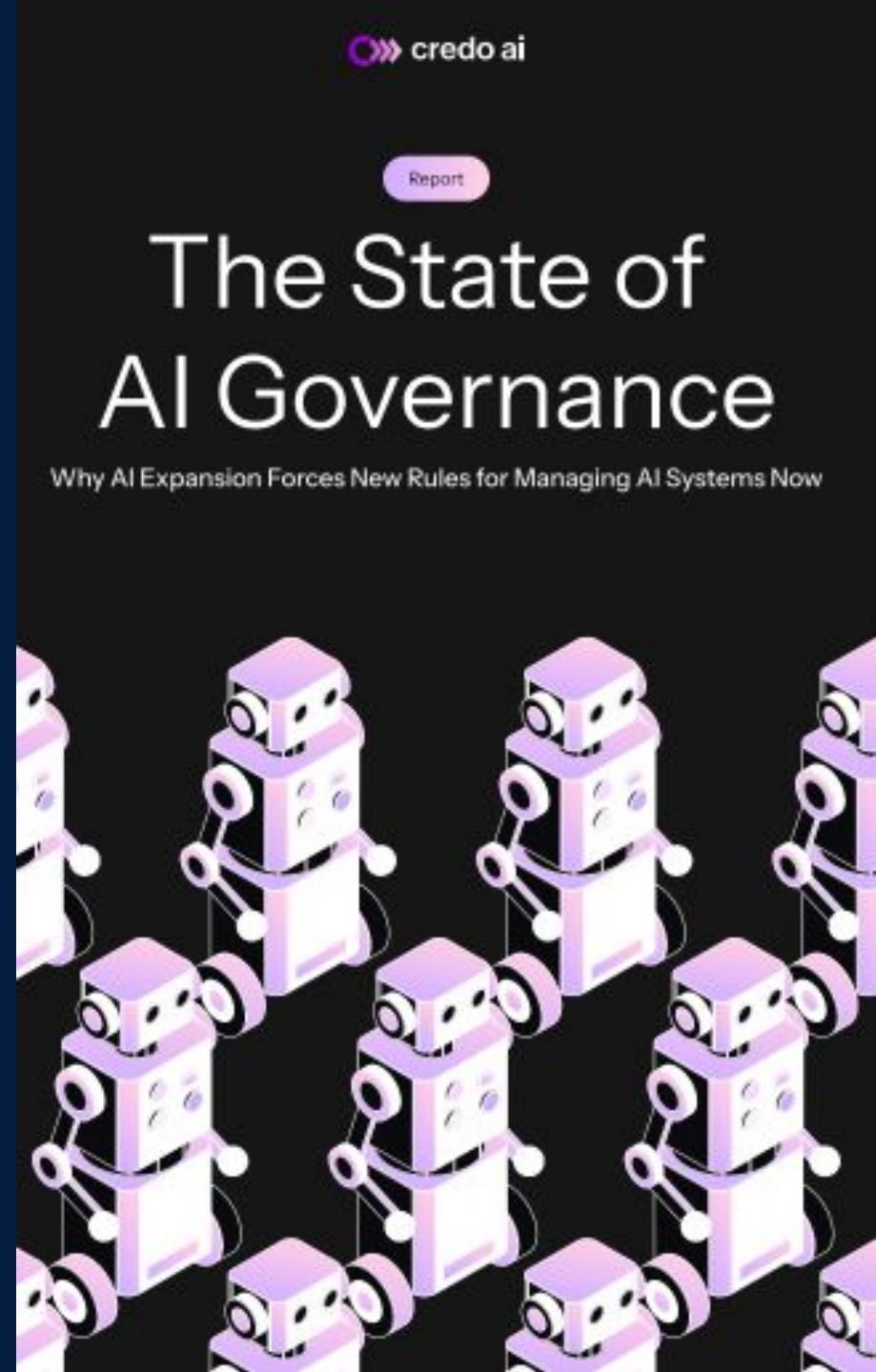
4% gobierna a escala

¿Qué hace ese 4%?

Convierte los marcos en operación: procesos definidos, estructuras de ownership y herramientas para desplegar IA a escala.

No principios en un PDF — un sistema que opera.

Fuentes: MIT NANDA (GenAI Divide 2025) · Credo AI, The State of AI Governance (pág. 17)



Marco ETO: extender el gobierno que ya tienes

No es una estructura nueva: es el gobierno corporativo extendido a la IA, en un circuito bidireccional.



Baja la dirección

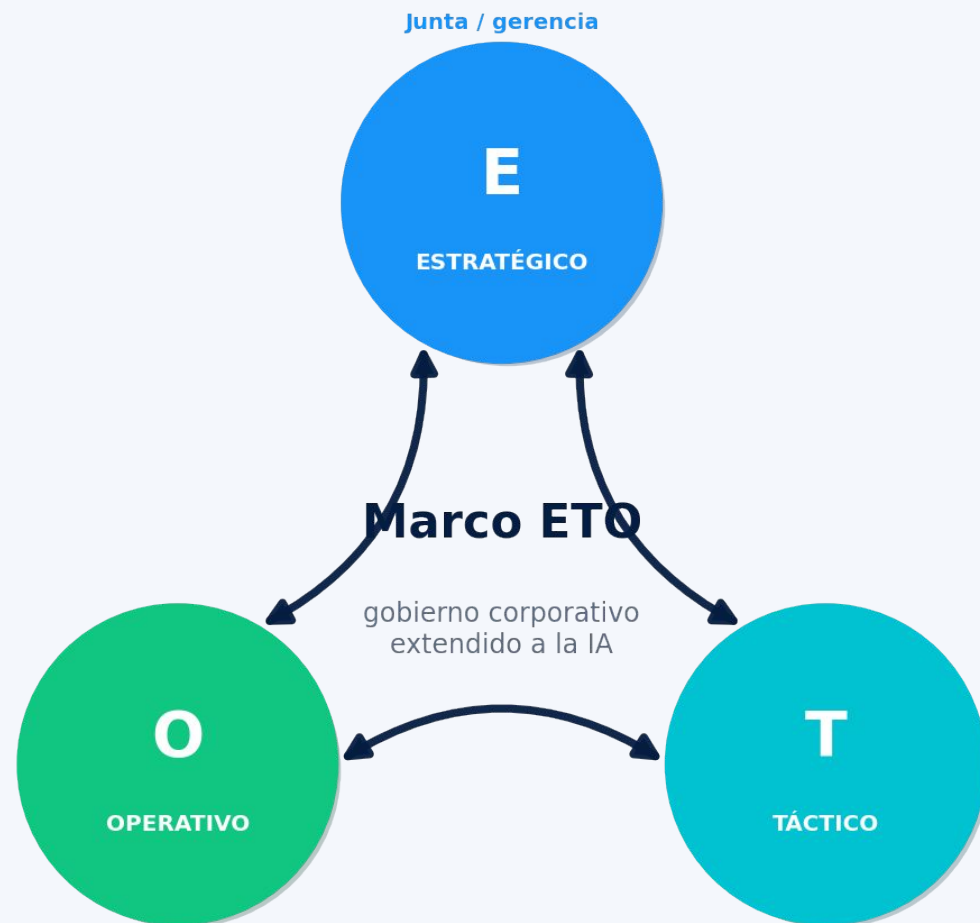
La alta dirección define propósito, principios y apetito de riesgo.



Sube el aprendizaje

La operación devuelve fricciones, casos límite e incidentes reales.

Un marco, cinco estándares: ISO 42001 · ISO 27001 · NIST AI RMF · NIST CSF 2.0 · RAI MM (Microsoft).



Etapa 1 · Estratégico

¿Para qué usamos IA y hasta dónde estamos dispuestos a llegar?

Quién responde: Junta, socios o gerencia general

QUÉ SE HACE

- Mandato escrito de la dirección.
- Propósito + apetito de riesgo declarado.
- Perímetro de cumplimiento.
- Responsable nombrado con autoridad real.



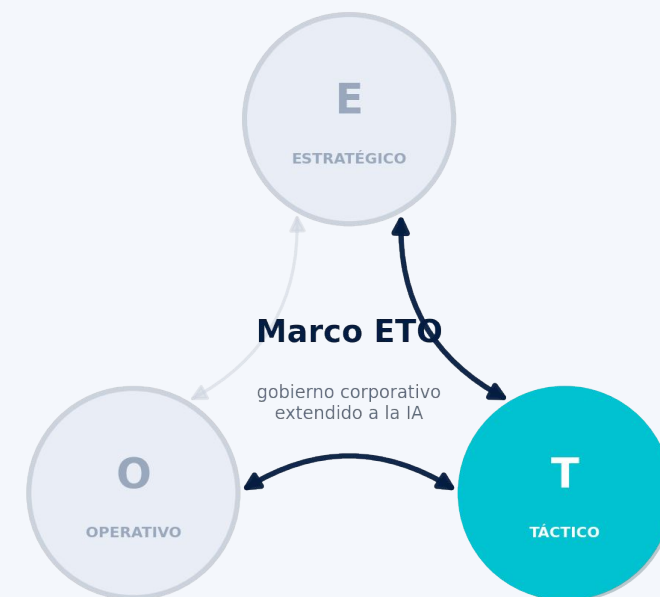
Etapa 2 · Táctico

¿Qué tenemos, qué riesgo tiene y con qué reglas lo manejamos?

Quién responde: Gerencias de área + responsable de gobernanza

QUÉ SE HACE

- Inventario de IA (incluida la shadow AI).
- Clasificación por riesgo regulatorio y de impacto de negocio.
- Política unificada de IA.
- Controles mínimos con dueño asignado y evidencia esperada: revisión humana de salidas, control de acceso, clasificación de la información, registro de decisiones y trazabilidad de versiones.
- Debida diligencia de proveedores: gobernanza de IA del proveedor, gestión y uso de los datos del cliente, seguridad y cumplimiento, y evidencia documental verificable
- Arquitectura de decisión.
- Plan de formación.



Etapa 3 · Operativo

¿Funciona en la práctica y qué aprendemos de ello?

Quién responde: Dueños de proceso, TI y proveedores

QUÉ SE HACE

- Evaluación de impacto y registro
- Monitoreo del sistema y del proveedor.
- Gestión de incidentes (IA + seguridad)
- Retroalimentación ascendente (pruebas, reunión trimestral, canal de reporte)
- Remedición de madurez + revisión por la dirección



Gobernar la IA trae ventajas competitivas

Bien hecha, la gobernanza es una palanca de crecimiento, ventaja competitiva y reputación — no un centro de costos.

Mejora tus proyecciones de negocio

+27%

de probabilidad de captar capital con prácticas ESG sólidas

59%

reporta impacto positivo por invertir operativamente en ESG

3×

más crecimiento en ingresos por colaborador en industrias más expuestas a la IA

Genera ventajas competitivas

ROI 2×

superior en proyectos de IA vs. empresas sin gobernanza responsable

+50%

en adopción tecnológica, cumplimiento de objetivos y aceptación de usuarios

Fortalece tu reputación

93%

de la Gen Z quiere que las empresas respalden sus compromisos ESG

Requisito

las grandes compañías exigen estándares ESG para contratar proveedores

Fuentes: PwC · Deloitte · PwC (Barómetro Global de Empleos en IA 2025) · IBM · Gartner · Porter Novelli · Observatorio de la Inversión ESG

En la era digital, **la confianza** es el activo más valioso de una organización.

Quien la construye —a través de una gobernanza para una IA responsable— gana la carrera de la adopción digital y percibe un verdadero retorno de la inversión: no solo en eficiencia, sino en reputación, talento y acceso a mercados.

En tu empresa ya se usa IA (aunque no lo sepas)



Pero ...¿Sabes en dónde estás y hacia dónde debes ir?

1. Diagnóstico de madurez

GRATUITO

Responde 30 preguntas y recibe un reporte inmediato con tu hoja de ruta para implementar IA Confiable.

Código de acceso: INNOVAITORS



2. Mentoría Estratégica

GRATUITO

- 10 cupos exclusivos
- 30 mins
- Asesoría sobre cómo implementar IA en tu empresa

