

La guía de decisión de la supervisión humana

Abreviaturas y acrónimos

Ningún conocimiento previo se da por supuesto. Los acrónimos de este documento:

IA : inteligencia artificial

UE : Unión Europea

¿Cuándo debe un ser humano permanecer al mando de una decisión tomada por inteligencia artificial (IA)? Esta guía responde en una página, con los tres modelos de control y las cinco pruebas de una supervisión real.

Los tres modelos de control

Humano en el circuito: la IA propone, una persona aprueba cada decisión antes de que surta efecto. Es el modelo para decisiones de alto impacto sobre personas. Humano sobre el circuito: la IA decide, una persona vigila el conjunto y puede intervenir. Es el modelo para volúmenes altos con riesgo medio. Humano al mando: una persona conserva la autoridad de detener el sistema entero en cualquier momento, sin pedir permiso. Este modelo no es opcional: acompaña a los otros dos siempre.

Cómo elegir

Pregunte una sola cosa: ¿qué pierde una persona concreta si la decisión sale mal? Si la respuesta toca crédito, salud, empleo, educación o servicios públicos, el Reglamento de IA de la UE exige supervisión humana efectiva para los sistemas de alto riesgo, y el modelo correcto es el humano en el circuito.

Las cinco pruebas de una supervisión real

Una supervisión que no pasa estas pruebas es decorativa. Uno: la persona entiende lo que el sistema hace y sus límites. Dos: tiene tiempo real para revisar, no tres segundos. Tres: tiene autoridad para rechazar la salida del sistema sin penalización. Cuatro: tiene acceso a la información que la decisión usó. Cinco: alguien mide si sus intervenciones ocurren de verdad. Un aprobador que aprueba todo no es supervisión: es una firma de goma.

El error más común

Instalar la supervisión al final, como un sello. La supervisión efectiva se diseña dentro del flujo: quién revisa, cuándo, con qué pantalla y con qué poder. Si el diseño no lo permite, el sistema no está gobernado.

El puente hacia la profundidad

Esta guía es el QUÉ. El CÓMO arquitectónico (los puntos de control, los patrones de diseño y las plantillas de revisión) vive en el AI Architecture Governance Blueprint de la capa de pago.

AI GOVERNANCE CHAIN™

Marco de gobernanza de la IA

© 2026 Mathieu K. Gouanou. Todos los derechos reservados.

Miembro del Consejo Asesor de Harvard Business Review.

Norma del documento: gobernanza de la IA centrada en el ser humano.