Edición de lectura | 2026-11-20

Autoridad para suspender una decisión de IA

Adrian Dunkley | 2026-11-15

El Consejo examina si la persona que revisa un pago puede detenerlo antes de la transferencia. Compara pérdidas simuladas con distintos niveles de personal y fallos compartidos. Los importes corresponden a una organización hipotética.

El Consejo examina si la persona que revisa un pago puede detenerlo antes de la transferencia. Compara pérdidas simuladas con distintos niveles de personal y fallos compartidos. Los importes corresponden a una organización hipotética.

Método y alcance

Un comité de riesgos puede poner a prueba el momento de la aprobación, los fallos compartidos y los límites de ejecución antes de ampliar el proyecto piloto. Pida al responsable del proceso que identifique la primera acción costosa de revertir. Un revisor autorizado debe comprobar sus premisas. Fije un plazo y suspenda la acción si el revisor no está disponible. Separe las verificaciones previas de las posteriores.

Resultados y ejemplo

En el análisis ilustrativo de Adrian Dunkley, la supervisión experta redujo la pérdida anual esperada a US$317.000. Al reducir a la mitad el tiempo del personal y la proporción de revisiones que verificaban premisas decisivas, y duplicar la probabilidad de un fallo compartido, la pérdida subió a US$447.000. La referencia humana fue US$431.000. El umbral superado en cerca del 1 % de los años simulados subió de US$1,42 millones a US$1,49 millones.

Uso y límites

Simule una caída del servicio de aprobación o de una fuente de datos del proveedor. Registre las acciones afectadas, el tiempo de detención y si se pudo deshacer la operación. Antes de elevar los límites de transacción, pida a la persona responsable que acepte el riesgo restante. Consulte el resumen del modelo y sus pruebas.

Referencias

Publicación original en inglés

Adrian Dunkley · Who can stop an AI decision?

Edición en inglés (PDF)