

Checklist para evaluar herramientas de IA

RECURSO DE TRABAJO

Una ficha de preevaluación para negocio, compras, privacidad, seguridad y cumplimiento. Sirve para decidir si se rechaza, se escala, se pilota o se aprueba con controles.

PUERTA DE ENTRADA

Primero, descarta lo que no debe avanzar

No compenses una señal crítica con una puntuación alta. Si aparece un bloqueo, detén la evaluación y escala a las funciones competentes.

SEÑALES DE BLOQUEO O ESCALADO INMEDIATO

- El caso de uso puede encajar en una práctica prohibida o en un sistema de alto riesgo.
- La herramienta decidirá o recomendará sobre empleo, crédito, salud, educación o acceso a servicios esenciales.
- Se prevé introducir datos sensibles, secretos o información de terceros sin base, control o autorización claros.
- El proveedor puede reutilizar datos o prompts para entrenar sin una opción contractual adecuada.
- La herramienta ejecutará acciones relevantes sin aprobación humana, trazabilidad o reversión.
- No hay contrato, identidad verificable del proveedor, canal de incidentes o posibilidad de borrar/exportar datos.

Flujo de decisión

01

Necesidad

Problema, propietario y alternativa sin IA

02

Riesgo

Personas, datos, impacto y clasificación

03

Evidencia

Proveedor, contrato, seguridad y pruebas

04

Piloto

Controles, métricas, supervisión y salida

05

Decisión

Aprobar, condicionar, escalar o rechazar

01 · NECESIDAD

Caso de uso y responsabilidad

Una herramienta no se evalúa en abstracto: se evalúa para un propósito, unas personas y un proceso concretos.

MARCA UNA OPCIÓN POR CONTROL

☐ SÍ ☐ PARCIAL ☐ NO ☐ N/A

01	El problema de negocio y el resultado esperado están descritos sin depender del nombre comercial de la herramienta.	☐	☐	☐	☐
02	Existe una persona propietaria del caso de uso y una función responsable de su validación.	☐	☐	☐	☐
03	Se ha considerado una alternativa sin IA o con menor exposición al riesgo.	☐	☐	☐	☐
04	Las personas usuarias, administradoras y potencialmente afectadas están identificadas.	☐	☐	☐	☐
05	Se conocen las decisiones o tareas que la salida apoyará, automatizará o condicionará.	☐	☐	☐	☐
06	Se ha definido qué ocurre si la herramienta falla, no está disponible o entrega una respuesta incorrecta.	☐	☐	☐	☐
07	El beneficio esperado tiene métricas verificables de calidad, tiempo, coste o reducción de riesgo.	☐	☐	☐	☐
08	El alcance excluye expresamente usos, datos y acciones que no están autorizados.	☐	☐	☐	☐
09	Hay presupuesto, recursos de implantación, formación y mantenimiento durante todo el ciclo de vida.	☐	☐	☐	☐
10	La evaluación tiene fecha, responsable y versión exacta de herramienta o modelo.	☐	☐	☐	☐
● Define una unidad de evaluación: herramienta + versión + caso de uso + tipo de datos + integración. Si cambia uno, revisa el riesgo.					

02 · PROVEEDOR

Contrato, servicio y cadena de suministro

La web comercial no basta. Pide evidencias contractuales y operativas que puedan conservarse.

MARCA UNA OPCIÓN POR CONTROL

☐ SÍ ☐ PARCIAL ☐ NO ☐ N/A

01	La entidad proveedora, jurisdicción, contacto y responsables del servicio están identificados.	☐	☐	☐	☐
02	Se dispone de contrato, términos aplicables a empresa y orden de prevalencia entre documentos.	☐	☐	☐	☐
03	Están documentados niveles de servicio, soporte, mantenimiento y cambios de funcionalidad o modelo.	☐	☐	☐	☐
04	Se conocen subencargados, subproveedores, modelos de terceros e infraestructura relevante.	☐	☐	☐	☐
05	El proveedor notifica incidentes, cambios materiales, discontinuación y alteraciones de condiciones.	☐	☐	☐	☐
06	Se puede exportar información, recuperar configuraciones y cerrar el servicio sin dependencia desproporcionada.	☐	☐	☐	☐
07	Las reglas sobre uso de entradas, salidas y metadatos para entrenamiento o mejora son aceptables y contractuales.	☐	☐	☐	☐
08	Se aclaran titularidad, licencia y restricciones sobre prompts, documentos, resultados y contenidos generados.	☐	☐	☐	☐
09	Existe compromiso de borrado, devolución y conservación limitada al finalizar el contrato.	☐	☐	☐	☐
10	Las declaraciones de seguridad, privacidad y cumplimiento están respaldadas por evidencia vigente y verificable.	☐	☐	☐	☐

● Conserva la versión de términos, anexos y evidencias usada para decidir. Un enlace vivo puede cambiar sin dejar rastro.

03 · DATOS

Protección de datos y confidencialidad

Evalúa por separado datos de entrada, metadatos, memoria, registros, resultados y datos que el proveedor genera o infiere.

MARCA UNA OPCIÓN POR CONTROL

☐ SÍ ☐ PARCIAL ☐ NO ☐ N/A

01	Se ha inventariado cada categoría de dato que entrará, se generará, almacenará o compartirá.	☐	☐	☐	☐
02	Se distingue entre datos personales, categorías especiales, confidenciales, secretos y datos de terceros.	☐	☐	☐	☐
03	La finalidad y la base jurídica del tratamiento están identificadas cuando existe dato personal.	☐	☐	☐	☐
04	Los roles RGPD están definidos y, cuando procede, existe un acuerdo de encargo conforme al artículo 28.	☐	☐	☐	☐
05	Se aplica minimización: solo entra el dato estrictamente necesario y se usan anonimización o seudonimización cuando encaja.	☐	☐	☐	☐
06	Se conocen ubicación, transferencias internacionales y salvaguardas aplicables a datos y soporte.	☐	☐	☐	☐
07	Retención, historial, memoria, copias y borrado se pueden configurar y verificar.	☐	☐	☐	☐
08	El uso de datos para entrenamiento, mejora, evaluación humana o telemetría está controlado y documentado.	☐	☐	☐	☐
09	Se ha valorado información a personas, ejercicio de derechos y exactitud de datos o inferencias.	☐	☐	☐	☐
10	Se ha analizado la necesidad de una evaluación de impacto y se ha escalado al DPD cuando corresponde.	☐	☐	☐	☐
11	Existen reglas para no introducir datos no autorizados y mecanismos técnicos que las apoyan.	☐	☐	☐	☐
12	El registro de actividades y la documentación de privacidad se actualizan antes de producción.	☐	☐	☐	☐

● No asumas que 'no entrenar' significa 'no conservar'. Verifica por separado entrenamiento, retención, logs, soporte y copias.

04 · SEGURIDAD

Arquitectura, acceso y respuesta

La seguridad depende del modo de uso: web aislada, API, complemento, agente con acciones o integración con sistemas internos.

MARCA UNA OPCIÓN POR CONTROL

☐ SÍ ☐ PARCIAL ☐ NO ☐ N/A

01	La arquitectura, flujos de información, integraciones, conectores y privilegios están diagramados.	☐	☐	☐	☐
02	Se admiten identidad corporativa, MFA, roles, mínimo privilegio y baja inmediata de accesos.	☐	☐	☐	☐
03	Existe separación entre entornos, organizaciones, proyectos y datos de distintos clientes o equipos.	☐	☐	☐	☐
04	El cifrado en tránsito y reposo, la gestión de claves y las copias están documentados.	☐	☐	☐	☐
05	Se registran accesos, cambios, prompts, acciones administrativas y eventos necesarios para investigar.	☐	☐	☐	☐
06	La organización puede exportar logs o recibirlos en su sistema de monitorización cuando el riesgo lo requiere.	☐	☐	☐	☐
07	Se evalúan prompt injection, fuga de instrucciones, extracción de datos y contenido malicioso en fuentes conectadas.	☐	☐	☐	☐
08	Las acciones automáticas tienen permisos limitados, confirmación humana y posibilidad de reversión.	☐	☐	☐	☐
09	Hay gestión de vulnerabilidades, pruebas independientes o certificaciones con alcance y vigencia claros.	☐	☐	☐	☐
10	El proveedor tiene canal y plazo de notificación de incidentes, así como responsables de coordinación.	☐	☐	☐	☐
11	Se han definido continuidad, indisponibilidad, degradación segura y procedimiento de retirada.	☐	☐	☐	☐
12	La configuración segura se valida de nuevo tras cambios relevantes de modelo, función o integración.	☐	☐	☐	☐

● Trata conectores y agentes como una ampliación de privilegios. La calidad del modelo no limita lo que una integración puede leer o ejecutar.

05 · CALIDAD

Riesgos propios de la IA

Mide el sistema con ejemplos representativos y criterios de aceptación definidos antes del piloto.

MARCA UNA OPCIÓN POR CONTROL

☐ SÍ ☐ PARCIAL ☐ NO ☐ N/A

01	Se han identificado errores plausibles, alucinaciones, omisiones y respuestas que parecen correctas sin serlo.	☐	☐	☐	☐
02	Existe un conjunto de pruebas representativo del idioma, sector, población y casos difíciles.	☐	☐	☐	☐
03	La calidad se mide con criterios verificables, no solo con impresiones de personas usuarias.	☐	☐	☐	☐
04	Se evalúan sesgos, efectos desiguales, accesibilidad y desempeño sobre grupos relevantes.	☐	☐	☐	☐
05	Las fuentes, citas y datos recuperados se pueden comprobar contra información autorizada.	☐	☐	☐	☐
06	Se conocen límites de contexto, actualización, memoria, aleatoriedad y cambios entre versiones.	☐	☐	☐	☐
07	Se prueban respuestas ante instrucciones conflictivas, entradas adversarias y contenido no fiable.	☐	☐	☐	☐
08	La persona supervisora recibe información suficiente y dispone de tiempo y autoridad reales para intervenir.	☐	☐	☐	☐
09	Se evita el sesgo de automatización: la interfaz y el proceso no empujan a aceptar la salida por defecto.	☐	☐	☐	☐
10	Se registran errores, anulaciones humanas, reclamaciones e incidentes para detectar degradación.	☐	☐	☐	☐
11	Hay umbrales de parada, revisión o retorno a proceso manual cuando la calidad cae.	☐	☐	☐	☐
12	Las pruebas se repiten tras cambios de modelo, datos, prompts de sistema, conectores o finalidad.	☐	☐	☐	☐

● Un piloto demuestra el comportamiento en un conjunto concreto; no prueba que la herramienta sea correcta para cualquier uso futuro.

06 · NORMATIVA

AI Act, transparencia y propiedad intelectual

La clasificación debe basarse en la finalidad y el contexto, no en la etiqueta comercial que use el proveedor.

MARCA UNA OPCIÓN POR CONTROL

☐ SÍ ☐ PARCIAL ☐ NO ☐ N/A

01	Se ha definido el papel de la organización en la cadena: proveedor, responsable del despliegue, importador, distribuidor u otro.	☐	☐	☐	☐
02	Se ha comprobado si el uso puede encajar en una práctica prohibida del artículo 5 del AI Act.	☐	☐	☐	☐
03	Se ha analizado si el sistema o su finalidad puede calificarse de alto riesgo y qué obligaciones activa.	☐	☐	☐	☐
04	Se han identificado deberes de transparencia para interacción con IA, emociones, biometría o contenido sintético/manipulado.	☐	☐	☐	☐
05	Las instrucciones del proveedor y la documentación necesaria están disponibles para las personas responsables.	☐	☐	☐	☐
06	Se han definido registros, información, supervisión humana y conservación exigibles para el caso concreto.	☐	☐	☐	☐
07	Los derechos sobre datos de entrada, materiales de terceros, prompts, resultados y modificaciones están claros.	☐	☐	☐	☐
08	Se aplican revisiones de copyright, marcas, imagen, voz, secretos empresariales y licencias antes de publicar o reutilizar.	☐	☐	☐	☐
09	La comunicación al público no atribuye a la IA cualidades, garantías o certificaciones que no están demostradas.	☐	☐	☐	☐
10	Se ha revisado normativa sectorial, laboral, consumo, accesibilidad y contratación que pueda añadir obligaciones.	☐	☐	☐	☐

● Si hay dudas sobre una práctica prohibida, alto riesgo o impacto en derechos, no cierres la ficha como 'riesgo bajo': escala la clasificación.

07 · OPERACIÓN

Piloto, controles y ciclo de vida

La aprobación debe incluir condiciones de uso, responsables, revisión y un mecanismo para retirar la herramienta.

MARCA UNA OPCIÓN POR CONTROL

☐ SÍ ☐ PARCIAL ☐ NO ☐ N/A

01	El piloto tiene alcance, duración, usuarios, datos permitidos y entorno separados de producción.	☐	☐	☐	☐
02	Se han acordado criterios de éxito, tolerancias, evidencias y condiciones de parada antes de probar.	☐	☐	☐	☐
03	Las personas participantes reciben protocolo, formación y canal de consulta adecuados al riesgo.	☐	☐	☐	☐
04	Se definen verificaciones obligatorias, fuentes aceptadas y aprobaciones humanas por tipo de salida.	☐	☐	☐	☐
05	La configuración aprobada, prompts de sistema, conectores y restricciones quedan versionados.	☐	☐	☐	☐
06	Se monitorizan uso real, calidad, coste, errores, incidentes y anulaciones humanas.	☐	☐	☐	☐
07	Hay responsable para revisar cambios de proveedor, modelo, términos, subencargados y funcionalidades.	☐	☐	☐	☐
08	La decisión de producción fija usos permitidos, prohibiciones, condiciones y fecha de próxima revisión.	☐	☐	☐	☐
09	Se documentan comunicación, soporte, escalado, respuesta a incidentes y retirada de acceso.	☐	☐	☐	☐
10	Existe plan de salida: exportación, borrado, sustitución, continuidad y cierre contractual.	☐	☐	☐	☐

☒ Aprobar una herramienta no aprueba todos sus usos. Publica el catálogo con finalidades y límites comprensibles para la plantilla.

Ficha de resolución

La puntuación ayuda a ordenar pendientes, pero no sustituye la valoración experta. Una señal de bloqueo prevalece sobre el total.

ÁREA	0 - 2	BLOQUEO	RESPONSABLE / EVIDENCIA
Necesidad y responsabilidad			
Proveedor y contrato			
Datos y privacidad			
Seguridad			
Calidad y riesgos de IA			
AI Act, transparencia y PI			
Operación y salida			

0 = SIN EVIDENCIA · 1 = PARCIAL / CONDICIONADO · 2 = VALIDADO

Decisión

☐ RECHAZADA ☐ ESCALADA ☐ PILOTO CONTROLADO ☐ APROBADA CON CONDICIONES ☐ APROBADA

CONDICIONES Y USOS AUTORIZADOS

DATOS Y USOS PROHIBIDOS

RESPONSABLES DE SUPERVISIÓN Y PARADA

PRÓXIMA REVISIÓN / DISPARADORES DE REVISIÓN

Fuentes de contraste

AI Act · Reglamento (UE) 2024/1689
eur-lex.europa.eu

AEPD · Guías y herramientas
aepd.es

ENISA · Artificial intelligence cybersecurity
enisa.europa.eu