SEGURIDAD DE IA · Nº 01 · V1.0

Cómo evaluar a un proveedor de IA: 40 preguntas y evidencia

> Lo que separa a un proveedor de inteligencia artificial que cumplirá de uno que no es la evidencia verificable, no los adjetivos: certificados con alcance, informes de terceros, benchmarks con método y fecha, contratos auditables. Esta guía para Hispanoamérica y España ordena la evaluación en diez preguntas de fondo y cuarenta concretas, cada una con la respuesta que descalifica y la evidencia exigible, con Chile como caso desarrollado y los equivalentes legales de México, Colombia, Perú, Argentina y España/UE en una tabla. Incluye el checklist descargable.

Por **David Pinto** — Fundador y arquitecto de la plataforma · **LinkedIn**
Revisión técnica: Innova y Cree

Publicado 17 ago 2026 · 36 min de lectura · PDF

Respuesta corta: evidencia, no adjetivos

Un proveedor de inteligencia artificial que cumplirá se distingue de uno que no por una sola cosa: puede mostrar evidencia verificable de cada afirmación que hace. «Seguro», «confiable» y «cumple la normativa» son adjetivos; un certificado cuya declaración de alcance nombra el servicio, un informe de un tercero con fecha, un benchmark con método e intervalo de confianza, un contrato con derecho de auditoría y un plan de salida en formatos abiertos son evidencia. Esta guía convierte esa idea en un instrumento de trabajo: diez preguntas de fondo, ordenadas por el

recorrido que hacen sus datos, y cuarenta preguntas concretas con dos columnas que rara vez aparecen en los checklists: la **respuesta inaceptable** —la que suena bien y no resiste verificación— y la **evidencia exigible**. Úsela completa en un proceso de compra, o solo las preguntas bloqueantes en una primera reunión.

> AVISO
>
> Esta guía es orientación técnica elaborada por un proveedor de sistemas de IA. No constituye asesoría legal: las referencias a la Ley 21.719, la Ley 21.663 y otras normas deben revisarse con asesoría jurídica para cada caso.

Por qué el checklist de siempre no sirve

Los cuestionarios de proveedores que circulan en compras y seguridad se escribieron para software convencional: cifrado, contraseñas, respaldos, continuidad. Siguen siendo necesarios, pero cubren una fracción pequeña de lo que un comprador regulado pregunta cuando el servicio incluye inteligencia artificial. Nuestra experiencia respondiendo cuestionarios de más de cincuenta controles a compradores regulados es consistente: cerca de un séptimo de las preguntas son ciberseguridad clásica; el resto es específico de IA. Qué modelo hay debajo y quién lo controla. Si el proveedor aprende de los datos del cliente. Si la inferencia en el extranjero es transferencia internacional. Cómo se defiende el sistema de una instrucción escondida en un documento. Qué puede hacer la IA por su cuenta. Cuánto se equivoca y quién lo supervisa. Un proveedor que llega con su cuestionario ISO de siempre deja la mayor parte del examen en blanco. Por eso esta guía usa una taxonomía propia de diez preguntas y las cruza con los marcos públicos que sí existen[6912].

Cómo usar las diez preguntas y las cuarenta concretas

Las diez preguntas de fondo siguen el recorrido de los datos: quién responde por el proveedor, qué modelo hay debajo, dónde están los datos, si se aprende de ellos, si alguien puede abusar del sistema, quién entra, qué hace la IA sola, cuánto se equivoca, qué pasa cuando falla y cómo termina la relación. Cada una se despliega en cuatro preguntas concretas con tres columnas: la pregunta tal como se le hace al proveedor, la respuesta inaceptable y la evidencia exigible. Tres reglas de uso. Primera: la columna de evidencia es la pregunta real; el enunciado es solo el título. Segunda: al menos una pregunta por bloque se verifica en sesión técnica contra el sistema, no contra el documento. Tercera: no se suman porcentajes sin leer las notas; cinco preguntas son bloqueantes por sí solas (4.1, 3.2, 7.2, 8.1 y 10.1). La figura 1 resume el mapa.

Figura 1. Las diez preguntas de fondo, ordenadas por el recorrido de los datos: de la gobernanza del proveedor a la salida del contrato. Cada tarjeta indica la evidencia que satisface la pregunta.

1. ¿Quién responde por ti, aparte de ti? (aseguramiento externo)

La primera pregunta de fondo no es técnica: es quién, además del propio proveedor, ha verificado lo que el proveedor afirma. La evidencia clásica son la certificación ISO/IEC 27001 con una declaración de alcance que nombre el servicio contratado[10], el informe SOC 2 Tipo II con su período de observación[11] y una prueba de intrusión independiente sobre la capa de IA. La respuesta inaceptable más frecuente es la certificación corporativa presentada como si cubriera el servicio, o el certificado del proveedor de nube presentado como propio. Cuando esas evidencias no existen —lo normal en un proveedor pequeño o joven— lo exigible cambia, no desaparece: controles equivalentes verificables en sesión técnica, derecho contractual de auditoría e inspección, evidencia bajo requerimiento y compromiso fechado de prueba de intrusión antes de producción. Y siempre, la lista nominal de subprocesadores con país y alcance de acceso, porque sin ella ninguna de las preguntas siguientes se puede evaluar[1].

PREGUNTA	RESPUESTA INACEPTABLE	EVIDENCIA EXIGIBLE
1.1 ¿Qué certificación o informe de tercero cubre exactamente el servicio que nos vas a prestar, y no solo a la empresa?	«Somos ISO 27001» sin una declaración de alcance que nombre el servicio; o el certificado del proveedor de nube presentado como propio.	Certificado ISO/IEC 27001 con declaración de alcance que nombre el servicio, o informe SOC 2 Tipo II con su período; si no existen, controles equivalentes por escrito y derecho de auditoría (pregunta 1.3). *Ref.: ISO/IEC 27001:2022 (NCh-ISO IEC 27001:2023); AICPA SOC 2.*
1.2 ¿Cuándo fue la última prueba de intrusión independiente sobre la capa de IA y qué alcance tuvo?	Ofrecer «red teaming interno» como sustituto de la prueba de tercero; o una prueba que solo cubrió el sitio corporativo.	Resumen ejecutivo del informe de un tercero (alcance, fecha, hallazgos cerrados) de menos de 12 meses sobre la versión del producto; si no existe, compromiso contractual con alcance y fecha antes de producción. *Ref.: OWASP Top 10 para aplicaciones LLM*

PREGUNTA	RESPUESTA INACEPTABLE	EVIDENCIA EXIGIBLE
		(2025/2026); NIST SP 800-218 (SSDF).
1.3 ¿Aceptas un derecho de auditoría e inspección directa, ejercido por nosotros o por un tercero que designemos?	«Solo compartimos nuestros certificados»; un acuerdo de confidencialidad redactado para impedir la inspección.	Cláusula de auditoría con plazo de aviso, alcance, frecuencia y costo; sala de evidencia bajo acuerdo de confidencialidad; sesión técnica con acceso a la configuración real. *Ref.: ISO/IEC 27001:2022 Anexo A 5.19-5.22 (relaciones con proveedores).*
1.4 ¿Quiénes son todos tus subprocesadores, en qué país operan y qué ven exactamente de nuestros datos?	«Usamos los mejores proveedores de nube» sin lista; una lista sin país ni alcance; «no hay terceros» cuando el modelo base es de un tercero.	Lista nominal de subprocesadores con país y alcance del acceso, aviso previo ante cambios y derecho de objeción. *Ref.: Ley 21.719, art. 15 bis y Título V (arts. 27-28); ISO/IEC 27001:2022 Anexo A 5.21.*

pregunta de fondo 1 — cuatro preguntas concretas con la respuesta que descalifica y la evidencia exigible.

2. ¿Qué modelo hay debajo y quién lo controla? (cadena de suministro)

Buena parte de los productos de IA que se venden a empresas son una capa de aplicación sobre el modelo de un tercero. No hay nada malo en ello; lo malo es ocultarlo. La pregunta de fondo es qué modelos base participan, con qué versión y licencia, si son de pesos abiertos o cerrados, y qué ocurre cuando el proveedor del modelo lo cambia o lo retira. La evidencia es un inventario de componentes de IA —modelos, bibliotecas de inferencia con versión, subprocesadores y fecha de corte— que el proveedor mantiene y entrega bajo acuerdo de confidencialidad; lo detallamos en el artículo sobre el inventario de componentes de IA. Dos respuestas descalifican:

«siempre usamos la última versión» —el sistema cambia sin que el contrato cambie— y un alias comercial presentado como versión fijada. Lo que se espera es versión fijada por configuración, integridad verificada de los artefactos que se ejecutan[8] y una regla de cambio de modelo con aviso previo y re-medición pareada contra la línea base antes de producción[612].

PREGUNTA	RESPUESTA INACEPTABLE	EVIDENCIA EXIGIBLE
2.1 ¿Qué modelos base usas (proveedor, nombre, versión, pesos abiertos o cerrados) y con qué licencia?	«IA propietaria» que oculta que es una capa sobre el modelo de un tercero; negativa a declararlo; un alias comercial presentado como versión.	Inventario de componentes de IA: modelos, bibliotecas de inferencia con versión, subprocesadores y fecha de corte, con compromiso de mantenerlo; entregable bajo acuerdo de confidencialidad. *Ref.: OWASP LLM03:2025 (cadena de suministro); NIST AI RMF (función MAP).*
2.2 ¿La versión del modelo está fijada? ¿Qué ocurre cuando el proveedor del modelo la cambia o la retira?	«Siempre usamos la última versión» (cambio silencioso); no conocer la fecha de retiro del modelo en uso.	Versión fijada por configuración; política de cambio de modelo con aviso previo y re-medición pareada contra la línea base antes de producción; capa de proveedor conmutable. *Ref.: NIST AI RMF (MEASURE, MANAGE); ISO/IEC 42001 (NCh-ISO IEC 42001:2024).*
2.3 ¿Cómo verificas la integridad de los artefactos que ejecutas (pesos abiertos, bibliotecas) y qué haces ante una vulnerabilidad publicada?	Descargas sin verificación de hash o firma; dependencias sin versión fijada; «actualizamos cuando podemos».	Verificación de integridad documentada; versiones exactas fijadas; procedimiento y plazo de actualización ante una vulnerabilidad publicada. *Ref.: NIST SP 800-218 (SSDF); OWASP LLM03:2025 y LLM04:2025.*

PREGUNTA	RESPUESTA INACEPTABLE	EVIDENCIA EXIGIBLE
2.4 ¿Los documentos que ingerimos salen a un tercero para indexarse (embeddings) o se procesan en tu infraestructura?	No saberlo; embeddings calculados por un tercero que no figura como subprocesador.	Declaración del lugar de cómputo de embeddings e indexación; si es un tercero, contrato, país y alcance; preferible: modelo de embeddings ejecutado en infraestructura propia. *Ref.: OWASP LLM08:2025 (vectores y embeddings); Ley 21.719, Título V.*

`pregunta de fondo 2 — cuatro preguntas concretas con la respuesta que descalifica y la evidencia exigible.`

3. ¿Dónde están mis datos y quién los toca? (residencia y transferencia)

«En la nube» no es una respuesta. La pregunta tiene dos partes: dónde se almacenan los datos y dónde se ejecuta la inferencia, porque pueden ser países distintos. Bajo la Ley 21.719, que entra en vigencia el 1 de diciembre de 2026, ejecutar el modelo fuera del país sobre datos personales es una operación de transferencia internacional aunque el almacenamiento sea local; el Título V (artículos 27 y 28) la permite hacia países con nivel adecuado de protección o amparada en cláusulas contractuales u otros instrumentos con garantías adecuadas[1]. La respuesta inaceptable es negar que exista transferencia porque «los datos se guardan en Chile». La evidencia es un diagrama de flujos con país por componente que identifique la única salida de datos del sistema y qué sale por ella, más la mitigación por diseño: solo fragmentos pertinentes, seudonimización, embeddings e índices en el país. Completan el bloque el aislamiento entre clientes demostrado con pruebas y el cifrado con modelo de llaves explícito.

> DATO

La Ley 21.719 fue publicada en el Diario Oficial el 13 de diciembre de 2024 y sus modificaciones a la Ley 19.628 entran en vigencia «el día primero del mes vigésimo cuarto posterior a la publicación», es decir, el 1 de diciembre de 2026; las multas por infracciones gravísimas llegan a 20.000 UTM (art. 35). Fuente: Biblioteca del Congreso Nacional, texto de la ley, consultado el 17 de agosto de 2026.

PREGUNTA	RESPUESTA INACEPTABLE	EVIDENCIA EXIGIBLE
3.1 ¿Dónde se almacenan los datos y dónde se ejecuta la inferencia? (pueden ser lugares distintos)	«En la nube»; responder solo por el almacenamiento y omitir dónde corre el modelo.	Diagrama de flujos con país por componente y la única salida de datos del sistema identificada; qué sale exactamente por esa salida. *Ref.: Ley 21.719, Título V (arts. 27-28).*
3.2 Si la inferencia ocurre fuera de Chile, ¿bajo qué mecanismo de transferencia internacional y con qué mitigación por diseño?	«No es transferencia porque los datos se guardan en Chile»; ampararse en un mecanismo que la ley no contempla.	Instrumento del art. 27 letra b) (cláusulas contractuales u otro instrumento con garantías adecuadas) y mitigación por diseño: solo fragmentos pertinentes, seudonimización, embeddings e índices en el país. *Ref.: Ley 21.719, arts. 27-28.*
3.3 ¿Cómo separas nuestros datos de los de otros clientes, unidades y sesiones?	«Es multiusuario y seguro» sin mecanismo; ninguna prueba de aislamiento realizada.	Aislamiento por cliente demostrado (instancia dedicada o segregación con pruebas); pruebas de fuga entre unidades incluidas en el alcance de la prueba de intrusión. *Ref.: ISO/IEC 27001:2022 Anexo A 8.22, 8.28; OWASP LLM02:2025.*
3.4 ¿Cifras en tránsito y en reposo, y quién controla las llaves?	«Todo está cifrado» sin modelo de llaves; llaves dentro del repositorio de código o en manos de personas.	TLS 1.2 o superior; cifrado en reposo con llaves fuera del código, rotación y revocación; llave de respaldos custodiada fuera del servidor. *Ref.: ISO/IEC 27001:2022 Anexo A 8.24; Ley 21.719, art. 14*

PREGUNTA	RESPUESTA INACEPTABLE	EVIDENCIA EXIGIBLE
		quinquies (seudonimización y cifrado).

pregunta de fondo 3 — cuatro preguntas concretas con la respuesta que descalifica y la evidencia exigible.

4. ¿Aprendes de mis datos, cuánto los guardas y cómo los borras?

Tres preguntas que parecen una sola y que los contratos genéricos no resuelven. Sobre el entrenamiento, la única respuesta aceptable es no entrenar por defecto y por contrato, con la cláusula extendida a los subprocesadores, incluido el proveedor del modelo base; los términos de los principales proveedores de modelos vía API declaran hoy que no entrenan con datos de clientes de API, y el comprador debe pedir el enlace y la fecha exacta[18][19]. Un opt-out en una consola no es una cláusula. Sobre la retención, lo exigible son plazos en días por tipo de dato, configurables, y la retención de registros del propio proveedor del modelo declarada con su fuente. Sobre el borrado, el procedimiento debe alcanzar índices vectoriales, derivados y todos los respaldos, con plazo máximo y certificado firmado, y servir también para una supresión puntual: el derecho de supresión de la Ley 21.719 no distingue entre la base productiva y el índice de búsqueda[1].

PREGUNTA	RESPUESTA INACEPTABLE	EVIDENCIA EXIGIBLE
4.1 ¿Usas nuestros datos para entrenar o mejorar modelos, tuyos o de tus proveedores?	«Solo con opt-out en la consola»; silencio sobre lo que hace el proveedor del modelo base.	Cláusula de no entrenamiento por defecto y por contrato, extendida a subprocesadores; enlace y fecha de los términos del proveedor del modelo que excluyen el entrenamiento. *Ref.: Ley 21.719, art. 15 bis (fin distinto prohibido); términos comerciales de los proveedores de modelos.*

PREGUNTA	RESPUESTA INACEPTABLE	EVIDENCIA EXIGIBLE
4.2 ¿Cuánto tiempo guardas prompts, salidas y documentos, y es configurable?	«Indefinidamente» o «no lo sabemos»; retención de registros del proveedor del modelo no declarada.	Plazos en días por tipo de dato, configurables; retención de registros del proveedor del modelo declarada con su fuente. *Ref.: Ley 21.719, art. 3 letra c) (proporcionalidad y plazo de conservación).*
4.3 ¿Cómo borras nuestros datos, incluidos índices vectoriales, derivados y respaldos, y qué certificado entregas?	«Se borran de la base» sin índices ni respaldos; sin certificado.	Procedimiento de eliminación que alcanza índices, derivados y todos los respaldos, con plazo máximo y certificado firmado; el mismo mecanismo para una supresión puntual. *Ref.: Ley 21.719, art. 4 (derecho de supresión) y art. 15 bis.*
4.4 ¿Qué envías exactamente al modelo: expedientes completos o fragmentos pertinentes?	«El documento entero» por defecto; sin minimización ni seudonimización.	Arquitectura de recuperación que envía solo los fragmentos pertinentes; seudonimización cuando corresponda; registro de lo enviado. *Ref.: Ley 21.719, art. 3 letra c) (proporcionalidad); NIST AI 600-1.*

pregunta de fondo 4 — cuatro preguntas concretas con la respuesta que descalifica y la evidencia exigible.

5. ¿Puede alguien hacer que tu IA haga algo indebido? (seguridad de la aplicación)

Aquí viven los riesgos que OWASP ordena en su Top 10 para aplicaciones con modelos de lenguaje —disponible en español en su edición 2025—: inyección de instrucciones, divulgación de información sensible, manejo inadecuado de salidas, debilidades en vectores y embeddings, consumo ilimitado[12][13][14]. La inyección indirecta merece atención especial: un documento, un correo o una planilla ingeridos pueden contener órdenes ocultas. La respuesta inaceptable es «el modelo tiene filtros». La evidencia es de diseño: el contenido ingerido se trata siempre como dato

y nunca como instrucción, el agente no tiene herramientas de ejecución que una instrucción inyectada pueda abusar, y existen pruebas hostiles documentadas con resultados. En la misma línea, las salidas se validan antes de mostrarse —las citas se contrastan por código contra la base— y los permisos se aplican en la recuperación, antes de que el contenido llegue al modelo, para que nadie vea a través de la IA lo que no vería en el sistema de archivos. Límites de tasa y red teaming documentado cierran el bloque[167].

PREGUNTA	RESPUESTA INACEPTABLE	EVIDENCIA EXIGIBLE
5.1 ¿Cómo tratas la inyección de instrucciones directa e indirecta (documentos, correos o planillas con órdenes ocultas)?	«El modelo tiene filtros»; «no nos ha pasado».	Diseño que trata el contenido ingerido como dato y nunca como instrucción; agente sin herramientas de ejecución; pruebas hostiles documentadas con resultados. *Ref.: OWASP LLM01:2025 (inyección de prompt).*
5.2 ¿Validas las salidas antes de mostrarlas o pasarlas a otro sistema (citas, enlaces, código, comandos)?	Salida cruda al usuario o a otro sistema; citas que nadie contrasta.	Verificación determinista de citas contra la base antes de mostrarlas; sanitización de salidas; ninguna acción se ejecuta a partir de texto generado sin validación. *Ref.: OWASP LLM05:2025 (manejo inadecuado de salidas).*
5.3 ¿Cómo evitas que el sistema revele información sensible o de otro usuario (permisos heredados en la recuperación)?	Un índice único sin control de acceso por documento; «el modelo sabe qué mostrar».	Filtro de permisos aplicado en la recuperación, antes de que el contenido llegue al modelo; pruebas de fuga entre usuarios y unidades. *Ref.: OWASP LLM02:2025 y LLM08:2025.*

PREGUNTA	RESPUESTA INACEPTABLE	EVIDENCIA EXIGIBLE
5.4 ¿Qué límites de consumo y de tasa aplicas, y existe red teaming de IA documentado?	Sin límites por usuario o llave; red teaming inexistente o solo verbal.	Límites de tasa y cuotas; informe interno de red teaming con categorías, fechas y hallazgos cerrados; guardarraíles configurables por el cliente. *Ref.: OWASP LLM10:2025 (consumo ilimitado); NIST AI 600-1; MITRE ATLAS.*

pregunta de fondo 5 — cuatro preguntas concretas con la respuesta que descalifica y la evidencia exigible.

6. ¿Quién entra, con qué, y cómo deja de entrar? (identidad y accesos)

Es el bloque más parecido a la seguridad clásica y, aun así, el que más olvidos acumula en sistemas con IA. Un usuario desvinculado que conserva acceso a la IA conserva acceso a todo lo que la IA indexa. Lo exigible es federación con el proveedor de identidad del cliente (OIDC o SAML) con la MFA del cliente, cuentas locales solo para soporte, nominadas y con expiración; una matriz de roles por acciones y niveles aplicada en el servidor en todos los puntos de entrada, no solo en la interfaz; secretos fuera del código, con alcance limitado, rotación y revocación; procesos automáticos con identidad de servicio propia; y desaprovisionamiento inmediato vía SSO con un plazo máximo para las cuentas locales, demostrable en la bitácora de accesos[10]. Las respuestas inaceptables son reconocibles: «administrador y usuario» como todo el modelo de roles, llaves de API en manos de usuarios finales, «cuando nos avisan» como plazo de baja.

PREGUNTA	RESPUESTA INACEPTABLE	EVIDENCIA EXIGIBLE
6.1 ¿Te integras con nuestro proveedor de	Cuentas locales con contraseña como	Federación OIDC o SAML con la MFA del proveedor de identidad del cliente;

PREGUNTA	RESPUESTA INACEPTABLE	EVIDENCIA EXIGIBLE
identidad (SSO con MFA)?	única opción; MFA «opcional».	cuentas locales solo para soporte, nominadas y con expiración. *Ref.: ISO/IEC 27001:2022 Anexo A 5.15-5.17, 8.5.*
6.2 ¿Cómo aplicas mínimo privilegio (roles por acciones) y dónde se aplica?	«Administrador y usuario»; permisos aplicados solo en la interfaz.	Matriz de roles por acciones y niveles aplicada en el servidor en todos los puntos de entrada; cuentas de administración de plataforma segregadas de la matriz del cliente. *Ref.: ISO/IEC 27001:2022 Anexo A 5.15, 5.18, 8.2, 8.3.*
6.3 ¿Cómo gestionas llaves y secretos (alcance, rotación, revocación) y quién los tiene?	Llaves de API en manos de usuarios finales; secretos en el repositorio; sin rotación.	Secretos fuera del código, con alcance limitado, rotación y revocación documentadas; procesos automáticos con identidad de servicio propia, nunca cuentas humanas. *Ref.: ISO/IEC 27001:2022 Anexo A 5.17, 8.24; OWASP LLM07:2025.*
6.4 ¿En cuánto tiempo deja de entrar un usuario desvinculado, y cómo lo demuestras?	«Cuando nos avisan», sin plazo; sin bitácora de accesos.	Baja inmediata vía SSO y plazo máximo para cuentas locales; revisión periódica de accesos; bitácora de accesos exportable. *Ref.: ISO/IEC 27001:2022 Anexo A 5.18, 8.15.*

pregunta de fondo 6 — cuatro preguntas concretas con la respuesta que descalifica y la evidencia exigible.

7. ¿Qué puede hacer la IA por su cuenta? (agencia)

Es la pregunta que más importancia ha ganado desde 2024: OWASP incluye la agencia excesiva en su Top 10 para aplicaciones LLM y publicó en diciembre de 2025 un Top 10 para aplicaciones agénticas, con el secuestro de objetivos y el mal uso de herramientas al frente[12][15]. La evaluación empieza por una lista cerrada: qué acciones

puede ejecutar la IA por su cuenta y qué sistemas alcanza, con declaración explícita de lo que no alcanza (por ejemplo, ninguna ruta a redes industriales u OT en infraestructura crítica). Sigue la compuerta humana: previa a cualquier acción con efecto fuera de la plataforma, por tipo de acción, registrada y no desactivable por configuración; una aprobación posterior no es una compuerta[1]. Y termina con dos controles operativos: la IA actúa con identidad de servicio propia, nunca con la cuenta de un usuario, y el cliente puede interrumpir el módulo de IA sin detener el resto de la plataforma. Un sistema que solo lee el corpus autorizado y redacta borradores cierra este bloque casi por diseño.

PREGUNTA	RESPUESTA INACEPTABLE	EVIDENCIA EXIGIBLE
7.1 Enumera las acciones que la IA puede ejecutar por su cuenta y los sistemas que alcanza.	«Puede hacer lo que le pidas»; una lista incompleta o que cambia sin aviso.	Lista cerrada de acciones y sistemas alcanzados; declaración explícita de lo que no alcanza (p. ej., ninguna ruta a redes industriales u OT); diagrama de puertos en sesión técnica. *Ref.: OWASP LLM06:2025 (agencia excesiva); OWASP Top 10 para aplicaciones agénticas 2026 (ASI02).*
7.2 ¿Existe una compuerta humana previa a cualquier acción con efecto fuera de la plataforma? ¿Se puede desactivar?	Aprobación posterior a la acción; compuerta configurable «para ir más rápido».	Compuerta humana previa, por tipo de acción, registrada en la bitácora y no desactivable por configuración; el revisor queda identificado. *Ref.: Ley 21.719, art. 8 bis (intervención humana); NIST AI RMF (GOVERN).*
7.3 ¿Con qué identidad actúa la IA cuando toca otros sistemas?	Con la cuenta del usuario o con una cuenta humana compartida.	Identidad de servicio propia, con permisos mínimos y acciones trazables a esa identidad. *Ref.: OWASP Top 10 para aplicaciones agénticas 2026 (ASI03); ISO/IEC 27001:2022 Anexo A 5.16.*
7.4 ¿Podemos interrumpir el módulo de IA de inmediato sin	«Hay que abrir un ticket»; una interrupción que	Interruptor del módulo de IA a disposición del cliente; comportamiento definido y probado ante la interrupción; el resto de la

PREGUNTA	RESPUESTA INACEPTABLE	EVIDENCIA EXIGIBLE
detener el resto de la plataforma?	deja acciones a medias.	plataforma sigue operando. *Ref.: NIST AI RMF (MANAGE); ISO/IEC 42001.*

pregunta de fondo 7 — cuatro preguntas concretas con la respuesta que descalifica y la evidencia exigible.

8. ¿Cuánto se equivoca, y cómo lo voy a saber? (exactitud y trazabilidad)

Ningún proveedor serio promete que su sistema no se equivoca; mide cuánto y lo publica. El estudio de Stanford RegLab sobre herramientas de investigación jurídica con IA lo demostró con datos: productos comercializados como libres de alucinaciones presentaron tasas de alucinación de 17 % y 33 % bajo una rúbrica estricta[17]. Lo exigible es un benchmark con rúbrica, tamaño de muestra, intervalo de confianza, pruebas hostiles y fecha, con compromiso de re-medición y la disposición a ejecutarlo frente al comité con sus propias preguntas. La segunda evidencia es la abstención: qué hace el sistema cuando el corpus no tiene la respuesta, medido como categoría propia. La tercera es la supervisión humana con revisor identificado y registro. La cuarta es la bitácora: consulta, salida, versión del modelo, usuario y hora, inmutable, con retención pactada y exportable al SIEM del cliente. Una bitácora sin la versión del modelo no permite reconstruir por qué el sistema respondió lo que respondió.

DATO

Stanford RegLab midió tasas de alucinación de 17 % (Lexis+ AI) y 33 % (Westlaw AI-Assisted Research) sobre más de 200 consultas jurídicas preregistradas, en productos comercializados como libres de alucinaciones. Fuente: Magesh et al., «Hallucination-Free? Assessing the Reliability of Leading AI Legal Research Tools»,

arXiv 2405.20362 (2024) y Journal of Empirical Legal Studies (2025), consultado el 17 de agosto de 2026.

PREGUNTA	RESPUESTA INACEPTABLE	EVIDENCIA EXIGIBLE
8.1 ¿Cuál es tu tasa de error medida, con qué método, sobre qué conjunto de consultas y cuándo?	Prometer que no se equivoca; una cifra sin método, sin intervalo de confianza y sin fecha.	Benchmark publicado con rúbrica, tamaño de muestra, intervalo de confianza, pruebas hostiles y fecha; compromiso de re-medición periódica; ejecutable frente al comité con sus propias preguntas. *Ref.: Stanford RegLab «Hallucination-Free?» (2024/2025); NIST AI 600-1 (confabulación).*
8.2 ¿Qué hace el sistema cuando el corpus no tiene la respuesta?	Responde siempre; sin abstención; sin nivel de confianza ni citas.	Abstención explícita («no está en las fuentes autorizadas») medida como categoría propia; citas verificables en cada salida. *Ref.: NIST AI 600-1; NIST AI RMF (MEASURE).*
8.3 ¿Quién supervisa las salidas y cómo queda registrado?	«El usuario es responsable», sin flujo de revisión ni revisor identificado.	Toda salida es un borrador con fuentes; flujo de revisión con revisor identificado y registro de quién aprobó y cuándo. *Ref.: Ley 21.719, art. 8 bis; ISO/IEC 42001.*
8.4 ¿Qué registra la bitácora (consulta, salida, versión de modelo, usuario, hora), cuánto se retiene y es exportable a nuestro SIEM?	Registros genéricos de aplicación; sin versión de modelo; bitácora editable.	Bitácora inmutable con encadenamiento verificable y campos mínimos definidos; retención pactada; exportación JSON/CSV o API para el SIEM del cliente. *Ref.: ISO/IEC 27001:2022 Anexo A 8.15-8.16; Ley 21.663, art. 9 (información para el reporte).*

pregunta de fondo 8 — cuatro preguntas concretas con la respuesta que descalifica y la evidencia exigible.

9. ¿Qué pasa cuando algo falla? (continuidad e incidentes)

En Chile este bloque dejó de ser una buena práctica para convertirse en un plazo legal. Las instituciones obligadas por la Ley 21.663 deben enviar al CSIRT Nacional una alerta temprana en un máximo de tres horas desde que conocen un incidente con impacto significativo, actualizarla en 72 horas —24 si un operador de importancia vital ve afectado su servicio— e informar en quince días[2]. Si el incidente ocurre en el servicio de IA de un proveedor, el plazo de aviso de ese proveedor tiene que caber dentro del plazo del cliente. Por eso «según el SLA de soporte» es inaceptable: lo exigible es una escalera de severidades con aviso inicial en horas «con lo que se sepa», informes preliminar y final con causa raíz, un contacto de seguridad nominado 24×7 con escalamiento —un correo de soporte comercial en horario hábil no lo satisface— y cooperación forense. Cierran el bloque el comportamiento definido si cae la API del modelo (capa de proveedor conmutable, RTO y RPO) y la única prueba que vale sobre los respaldos: la última restauración ejecutada, con fecha y verificación.

> DATO
>
> La Ley 21.663 fija plazos de reporte al CSIRT Nacional (art. 9): alerta temprana en un máximo de tres horas desde que se conoce el incidente, actualización en 72 horas (24 horas para un operador de importancia vital con su servicio afectado) e informe final en quince días corridos. Fuente: Biblioteca del Congreso Nacional, texto de la ley, consultado el 17 de agosto de 2026.

PREGUNTA	RESPUESTA INACEPTABLE	EVIDENCIA EXIGIBLE
9.1 ¿En cuántas horas nos avisas de un incidente, y cabe ese plazo dentro de nuestra obligación legal de reporte?	«Según el SLA de soporte» (días); avisar solo cuando esté el análisis completo.	Escalera de severidades con aviso inicial en horas «con lo que se sepa», informe preliminar y final con causa raíz; el plazo cabe dentro de las 3 horas de alerta temprana del art. 9 de la Ley 21.663 si el cliente está obligado. *Ref.: Ley 21.663, art. 9; Ley 21.719, art. 14 sexies.*
9.2 ¿Quién es tu contacto de seguridad 24×7 y cómo escala?	Un correo de soporte comercial en horario hábil.	Contacto de seguridad nominado, con escalamiento de dos niveles, en anexo a la firma; cooperación forense y preservación de evidencia comprometidas. *Ref.: Ley 21.663, art. 8 letra i) (delegado de ciberseguridad como contraparte); Ley 21.459.*
9.3 ¿Qué pasa si cae la API del modelo o su proveedor lo retira?	«El servicio se detiene»; dependencia de un único proveedor sin plan.	Comportamiento degradado definido; capa de proveedor conmutable; SLA con RTO y RPO. *Ref.: ISO/IEC 27001:2022 Anexo A 5.29-5.30; NIST CSF 2.0.*
9.4 ¿Cuándo fue la última restauración de respaldo que ejecutaste y qué verificaste después?	«Tenemos respaldos diarios» sin ninguna restauración probada; llave de los respaldos en el mismo servidor.	Registro de restauración con fecha y verificación (conteos, integridad de la bitácora); llave privada custodiada fuera del servidor; RPO y RTO en dos columnas: medido hoy y comprometido por contrato. *Ref.: ISO/IEC 27001:2022 Anexo A 8.13; Ley 21.719, art. 14 quinquies letra c).*

pregunta de fondo 9 — cuatro preguntas concretas con la respuesta que descalifica y la evidencia exigible.

10. ¿Cómo me voy, y de quién es lo que produce? (salida y legal)

La pregunta de salida responde al riesgo de dependencia de cualquier proveedor, grande o pequeño, y se contesta con arquitectura y contrato, no con organigrama. Lo exigible es un anexo de portabilidad: datos con su esquema, documentos en formato nativo descifrados, bitácora íntegra verificable y configuraciones, más asistencia de transición y, cuando el riesgo lo justifica, depósito de código fuente con eventos de liberación. «Pueden descargar desde la interfaz» es una respuesta inaceptable. En lo legal, tres puntos: los resultados del sistema son del cliente y el proveedor indemniza por propiedad intelectual de terceros[5]; el cumplimiento se ancla en leyes que existen —21.719, 21.663, 21.459, 20.393 y la norma sectorial que aplique, como el Capítulo 20-7 de la CMF en banca— con compromiso de adecuación futura, no en una «ley chilena de IA» que todavía no se ha promulgado[3420]; y el usuario sabe que interactúa con IA, con evaluación de sesgo e impacto cuando la decisión incide en personas naturales[1].

PREGUNTA	RESPUESTA INACEPTABLE	EVIDENCIA EXIGIBLE
10.1 ¿Cómo nos vamos: qué entregas, en qué formatos y con qué asistencia?	Exportación parcial o en formato propietario; «pueden descargar desde la interfaz».	Anexo de portabilidad: datos con su esquema, documentos en formato nativo descifrados, bitácora íntegra verificable, configuraciones; asistencia de transición; depósito de código fuente (escrow) con eventos de liberación cuando aplique. *Ref.: Ley 21.719, art. 4 (portabilidad) y art. 15 bis.*
10.2 ¿De quién son los resultados que produce el sistema y qué indemnidad ofreces por propiedad intelectual?	El proveedor se reserva derechos sobre salidas o datos; sin indemnidad.	Cláusula de titularidad de los resultados en favor del cliente; indemnidad por infracción de propiedad intelectual de terceros; términos del proveedor del modelo coherentes con lo anterior. *Ref.: Ley 17.336; términos comerciales del proveedor del modelo.*

PREGUNTA	RESPUESTA INACEPTABLE	EVIDENCIA EXIGIBLE
10.3 ¿A qué leyes vigentes anclas tu cumplimiento y cómo te adaptarás a las que vengan?	Invocar una «ley chilena de IA» que no existe; copiar una plantilla extranjera (RGPD/AI Act) sin adaptar.	Cláusula anclada en las leyes 21.719, 21.663, 21.459 y 20.393 y en la norma sectorial que corresponda (p. ej. RAN 20-7 de la CMF en banca), con compromiso de adecuación futura. *Ref.: Leyes 21.719, 21.663, 21.459, 20.393; RAN 20-7 (CMF).*
10.4 ¿El usuario sabe que interactúa con IA, y evalúas sesgo cuando la decisión incide en personas?	IA sin identificar; «no aplica» sin justificación.	Contenido generado identificado, con nivel de confianza y citas; evaluación de sesgo y de impacto cuando el caso de uso incide en personas naturales, con revisión humana. *Ref.: Ley 21.719, art. 8 bis y art. 15 ter; NIST AI RMF; ISO/IEC 42001.*

pregunta de fondo 10 — cuatro preguntas concretas con la respuesta que descalifica y la evidencia exigible.

Marco legal por país: la pregunta no cambia, la norma sí

Las diez preguntas valen igual en Santiago, Ciudad de México, Bogotá, Lima, Buenos Aires o Madrid. Lo que cambia es la norma que las vuelve exigibles y el plazo que fija. Esta tabla reúne, por jurisdicción, las tres familias de reglas que más pesan en un contrato de IA: datos personales y transferencia internacional (preguntas 3 y 4), ciberseguridad e incidentes (pregunta 9) y regulación específica de IA (preguntas 7 y 10). Consultada el 18 de agosto de 2026.

PAÍS	DATOS PERSONALES Y TRANSFERENCIA	CIBERSEGURIDAD E INCIDENTES	REGULACIÓN ESPECÍFICA DE IA
Chile	Ley 21.719 (D.O. 13 dic 2024, vigencia 1 dic 2026): encargado de tratamiento, transferencias	Ley 21.663, Marco de Ciberseguridad (D.O. 8 abr 2024): ANCI, deber de reportar incidentes con plazos escalonados	Sin ley vigente; proyectos en tramitación legislativa

PAÍS	DATOS PERSONALES Y TRANSFERENCIA	CIBERSEGURIDAD E INCIDENTES	REGULACIÓN ESPECÍFICA DE IA
	internacionales con garantías, decisiones automatizadas; crea la Agencia de Protección de Datos[1]	y obligaciones para operadores de importancia vital[2]	
México	Ley Federal de Protección de Datos Personales en Posesión de los Particulares, texto vigente compilado por la Cámara de Diputados (última reforma D.O.F. 14 nov 2025): encargado, transferencias, derechos ARCO [22]	Sin ley general; obligaciones sectoriales del supervisor que corresponda (financiero, salud, telecomunicaciones)	Sin ley específica; iniciativas en tramitación
Colombia	Ley 1581 de 2012 y su decreto reglamentario; autoridad de control: Superintendencia de Industria y Comercio [23]	Normativa sectorial del supervisor (para entidades vigiladas) y lineamientos de gobierno digital	Política pública de IA por documento CONPES; sin ley específica
Perú	Ley 29733 de Protección de Datos Personales y su reglamento (actualizado en 2025)	Sin ley general; normativa sectorial del supervisor	Ley 31814 que promueve el uso de la inteligencia artificial, y su reglamento (2025)
Argentina	Ley 25.326 de Protección de los Datos Personales; autoridad: AAIP [24]	Normativa sectorial del supervisor y lineamientos de la Administración Pública	Sin ley; recomendaciones y proyectos
España y UE	Reglamento (UE) 2016/679 (RGPD): encargado, transferencias internacionales,	Directiva (UE) 2022/2555 (NIS2), traspuesta por cada Estado miembro [26]	Reglamento (UE) 2024/1689 de inteligencia artificial, con calendario de

PAÍS	DATOS PERSONALES Y TRANSFERENCIA	CIBERSEGURIDAD E INCIDENTES	REGULACIÓN ESPECÍFICA DE IA
	decisiones automatizadas[25]		aplicación escalonado[27]

marco legal aplicable por jurisdicción a las preguntas de datos, incidentes y regulación de IA.

> **AVISO**
>
> Este cuadro orienta la conversación con un proveedor; no es asesoría legal y la vigencia cambia. Antes de redactar una cláusula, verifique el texto en la fuente primaria de su país y con su asesor. Verificamos el contenido de las fuentes de Chile, México, Colombia, Argentina y la Unión Europea desde nuestra propia infraestructura; para Perú no pudimos confirmar el enlace oficial y citamos las normas por su número, sin URL.

Tres reglas prácticas que se derivan de la tabla:

1. **Fije el plazo más corto que le aplique.** Si su empresa opera en varios países, el contrato debe obligar al proveedor al plazo de notificación más exigente de todos, no al del país donde está el servidor.
2. **La inferencia también es transferencia.** Si el modelo corre en otro país, sus datos salen aunque el almacenamiento sea local: eso activa las reglas de transferencia internacional en todas las jurisdicciones de la tabla.
3. **La ausencia de ley de IA no es ausencia de obligaciones.** Donde no hay norma específica, siguen aplicando datos personales, consumidor, propiedad intelectual y responsabilidad civil; y la norma llega después, así que el contrato debe incluir el deber de adecuarse a la que entre en vigor.

Cómo puntuar sin engañarse

El error más común al usar un checklist es sumar los SI y dividir. Tres correcciones evitan la mayor parte de los engaños. Primera: el criterio de aceptación es la pregunta real. «¿Tienen ISO 27001?» tiene como criterio «con una declaración de alcance que nombre el servicio»; sin esa condición, un SI vale poco. Segunda: «SI con evidencia» y «SI» no son lo mismo. Todo SI que llegue sin la evidencia de la columna correspondiente se registra como PARCIAL y se verifica en sesión técnica; en nuestra experiencia como proveedor, cada afirmación de un cuestionario debe auditarse contra el sistema real antes de firmarla, y cuando no coincide se corrige el sistema, no el texto. Tercera: el NO honesto vale más que el SI sin evidencia. Un proveedor que responde NO a una certificación que no tiene, y lo acompaña de controles equivalentes verificables, derecho de auditoría directa y un compromiso fechado, está mostrando madurez; el que responde SI a todo está pidiendo que se le crea. Un comité técnico serio desconfía del segundo.

Qué es razonable exigir a un proveedor pequeño (y qué no)

Exigir lo mismo a un proveedor global y a uno especializado de menor tamaño produce dos resultados igual de malos: descartar al mejor para el caso de uso o aceptar promesas que nadie verificará. La distinción útil es entre evidencias que cuestan años y evidencias que cuestan semanas. ISO/IEC 27001 con alcance del servicio y SOC 2 Tipo II son deseables y pueden pactarse como hoja de ruta contractual, pero exigirlas como condición de entrada deja fuera a casi todo proveedor especializado. Lo que sí es razonable exigir a cualquiera, sea cual sea su tamaño: prueba de intrusión de tercero sobre la versión evaluada del producto, renovada cada doce meses; inventario de componentes de IA; derecho de auditoría e inspección; tasa de error medida con método; compuerta humana no desactivable; cláusula de no entrenamiento extendida a subprocesadores; plazos de aviso de incidentes en horas; y plan de salida en formatos abiertos. Ninguna de esas

evidencias depende del tamaño de la organización; todas dependen de cómo está construido el sistema y de qué está dispuesto a firmar el proveedor.

Cómo respondemos nosotros

Por transparencia, y sin cifras ni nombres de ningún cliente, esto es lo que hoy podemos mostrar. Publicamos un benchmark de fiabilidad con su método: 278 consultas contra un sistema en producción bajo la rúbrica de Stanford RegLab, tasas de exactitud y de alucinación con sus intervalos de confianza, 0 fabricaciones en 120 pruebas hostiles y re-medición trimestral; la lectura honesta no es «cero», sino la cota superior con 95 % de confianza que la página publica[21]. Fijamos la versión del modelo por configuración y no la cambiamos sin re-medir contra la línea base con prueba pareada; la regla está en nuestra política de IA responsable. La supervisión humana es de diseño: toda salida es un borrador con citas verificadas por código, el sistema se abstiene cuando el corpus no tiene la respuesta y la bitácora es inmutable con encadenamiento verificable. El plan de salida entrega todo en formatos abiertos; el inventario de componentes de IA está disponible bajo solicitud y acuerdo de confidencialidad. Lo que aún no tenemos como evidencia de tercero: certificación ISO/IEC 27001 con alcance del servicio e informe SOC 2 Tipo II. Lo mitigamos con controles equivalentes verificables en sesión técnica, derecho contractual de auditoría e inspección directa, evidencia bajo requerimiento y prueba de intrusión independiente sobre la versión genérica del producto antes de producción en entornos regulados.

DATO

Nuestro benchmark público mide 278 consultas (medición del 2026-07-10, motor v2) con 84,2 % de respuestas exactas y 2,5 % de alucinación bajo la rúbrica de Stanford RegLab, intervalos de confianza de Wilson al 95 %, y 0 fabricaciones en 120 pruebas

hostiles. Fuente: innovaycree.com/empresas/benchmark/, consultado el 17 de agosto de 2026.

El checklist descargable

Las cuarenta preguntas están disponibles como planilla para enviarlas al proveedor con las columnas de respuesta vacías: número, pregunta de fondo, pregunta, respuesta inaceptable, evidencia exigible, referencia normativa, respuesta del proveedor (SI, NO, PARCIAL, N/A), evidencia entregada y nota. La hoja «Cómo usar» resume las reglas de puntuación de esta guía y la hoja «Licencia» indica los términos: CC BY-ND 4.0, con crédito a Innova y Cree y a la URL de este artículo, sin obras derivadas; si necesita una adaptación a otro sector o jurisdicción, escríbanos. Las mismas preguntas están en CSV, para importarlas a la herramienta de compras de su organización, y en Markdown, para versionarlas en un repositorio. Todas nuestras plantillas se reúnen en /recursos/plantillas/. Cuando actualicemos las preguntas —por cambios normativos, por una nueva edición de OWASP o por lo que aprendamos respondiendo cuestionarios— la versión y el historial quedarán registrados en esta página.

PLANTILLA DESCARGABLE

Checklist de evaluación de proveedores de IA (v1.0, XLSX)

Licencia CC BY-ND 4.0 · descargar

PLANTILLA DESCARGABLE

Checklist de evaluación de proveedores de IA (v1.0, CSV)

Licencia CC BY-ND 4.0 · descargar

PLANTILLA DESCARGABLE

Checklist de evaluación de proveedores de IA (v1.0, Markdown)

Si está evaluando un proveedor de IA para un entorno regulado y quiere contrastar sus respuestas con las nuestras, en la misma mesa y contra el sistema real: Diagnóstico sin costo.

1. Ley 21.719, BCN: art. primero transitorio (vigencia), arts. 3, 4, 8 bis, 14 quinquies, 14 sexies, 15 bis, 15 ter, 27-28 y 35. Consultado el 17 de agosto de 2026. ↩↩↩↩↩↩
2. Ley 21.663, BCN: art. 8 (deberes de los operadores de importancia vital) y art. 9 (deber de reportar). Consultado el 17 de agosto de 2026. ↩↩
3. Ley 21.459 sobre delitos informáticos, BCN. Consultado el 17 de agosto de 2026. ↩
4. Ley 20.393 sobre responsabilidad penal de las personas jurídicas, BCN. Consultado el 17 de agosto de 2026. ↩
5. Ley 17.336 sobre propiedad intelectual, BCN. Consultado el 17 de agosto de 2026. ↩
6. NIST, AI Risk Management Framework 1.0 (AI 100-1), enero de 2023. Consultado el 17 de agosto de 2026. ↩↩
7. NIST AI 600-1, Generative Artificial Intelligence Profile, julio de 2024. Consultado el 17 de agosto de 2026. ↩
8. NIST SP 800-218, Secure Software Development Framework v1.1. Consultado el 17 de agosto de 2026. ↩
9. ISO/IEC 42001:2023; adopción chilena NCh-ISO IEC 42001:2024 (INN). Consultado el 17 de agosto de 2026. ↩

10. ISO/IEC 27001:2022; adopción chilena NCh-ISO IEC 27001:2023 (INN). Consultado el 17 de agosto de 2026. ↩↩

11. AICPA & CIMA, SOC 2. Consultado el 17 de agosto de 2026. ↩

12. OWASP Top 10 for LLM Applications 2025 (LLM01-LLM10). Consultado el 17 de agosto de 2026. ↩↩↩↩

13. OWASP, Top 10 2025 de riesgos y mitigaciones para LLMs y aplicaciones de IA generativa (traducción al español). Consultado el 17 de agosto de 2026. ↩

14. OWASP GenAI LLM Top 10 — 2026, edición vigente (3-4 de agosto de 2026). Consultado el 17 de agosto de 2026. ↩

15. OWASP Top 10 for Agentic Applications for 2026 (ASI01-ASI10), diciembre de 2025. Consultado el 17 de agosto de 2026. ↩

16. MITRE ATLAS. Consultado el 17 de agosto de 2026. ↩

17. Magesh et al., «Hallucination-Free?», arXiv 2405.20362 (2024); J. Empirical Legal Studies (2025). Consultado el 17 de agosto de 2026. ↩

18. OpenAI, «Your data»: datos enviados a la API no se usan para entrenar salvo opt-in explícito. Consultado el 17 de agosto de 2026. ↩

19. Anthropic, Commercial Terms of Service: «Anthropic may not train models on Customer Content from Services». Consultado el 17 de agosto de 2026. ↩

20. CMF, RAN Capítulo 20-7, Externalización de servicios. Consultado el 17 de agosto de 2026. ↩

21. Innova y Cree, Benchmark anti-alucinación, medición de julio de 2026, re-medición trimestral. Consultado el 17 de agosto de 2026. ↩

Puntos clave

01 La evidencia verificable, no el adjetivo, es lo que separa a un proveedor que cumplirá de uno que no: certificado con alcance, informe de tercero, benchmark con método y fecha, contrato auditable.

02 De lo que exige un comprador regulado, solo una fracción es ciberseguridad clásica; el resto es específico de IA: cadena de modelos, no entrenamiento, transferencia por inferencia, inyección, agencia, exactitud medida y supervisión humana.

03 El criterio de aceptación es la pregunta real: «SI con evidencia» y «SI» no valen lo mismo, y un NO honesto con controles equivalentes y derecho de auditoría vale más que un SI sin evidencia.

04 La inferencia fuera del país es transferencia internacional bajo la Ley 21.719, y el plazo de aviso de incidentes del proveedor debe caber dentro de las tres horas de alerta temprana de la Ley 21.663 cuando el cliente está obligado.

05 A un proveedor pequeño no se le exige ISO 27001 ni SOC 2 como condición; se le exige prueba de intrusión de tercero, inventario de componentes, derecho de auditoría y un plan de salida en formatos abiertos.

Preguntas frecuentes

¿Un proveedor de IA sin ISO 27001 ni SOC 2 queda descartado?

¿Qué pregunta hago primero si solo tengo diez minutos?

Si los datos se guardan en Chile pero el modelo corre fuera, ¿hay transferencia internacional?

¿Cómo se relaciona la Ley 21.663 con un proveedor de IA?

¿Qué significa «respuesta inaceptable» en el checklist?

Referencias

[1] Ley 21.719 (Chile) — Regula la protección y el tratamiento de los datos personales y crea la Agencia de Protección de Datos Personales (D.O. 13 dic 2024; vigencia 1 dic 2026). https://www.bcn.cl/leychile/navegar?idNorma=1209272. Consultado el 17 de agosto de 2026. LEY

[2] Ley 21.663 (Chile) — Ley Marco de Ciberseguridad (D.O. 8 abr 2024). https://www.bcn.cl/leychile/navegar?idNorma=1202434. Consultado el 17 de agosto de 2026. LEY

[3] Ley 21.459 (Chile) — Establece normas sobre delitos informáticos. https://www.bcn.cl/leychile/navegar?idNorma=1177743. Consultado el 17 de agosto de 2026. LEY

[4] Ley 20.393 (Chile) — Responsabilidad penal de las personas jurídicas. https://www.bcn.cl/leychile/navegar?idNorma=1008668. Consultado el 17 de agosto de 2026. LEY

[5] Ley 17.336 (Chile) — Propiedad intelectual. https://www.bcn.cl/leychile/navegar?idNorma=28933. Consultado el 17 de agosto de 2026. LEY

[6] NIST AI Risk Management Framework 1.0 (AI 100-1), enero de 2023 — funciones GOVERN, MAP, MEASURE, MANAGE. https://www.nist.gov/itl/ai-risk-management-framework. Consultado el 17 de agosto de 2026. NORMA

[7] NIST AI 600-1 — Artificial Intelligence Risk Management Framework: Generative Artificial Intelligence Profile, julio de 2024. https://nvlpubs.nist.gov/nistpubs/ai/NIST.AI.600-1.pdf. Consultado el 17 de agosto de 2026. NORMA

[8] NIST SP 800-218 — Secure Software Development Framework (SSDF) v1.1. https://csrc.nist.gov/pubs/sp/800/218/final. Consultado el 17 de agosto de 2026. NORMA

[9] ISO/IEC 42001:2023 — Sistemas de gestión de inteligencia artificial (adopción chilena NCh-ISO IEC 42001:2024, INN, 2024). https://ecommerce.inn.cl/nch-iso-iec-42001202489243. Consultado el 17 de agosto de 2026. NORMA

[10] ISO/IEC 27001:2022 — Sistemas de gestión de la seguridad de la información (adopción chilena NCh-ISO IEC 27001:2023, INN, publicada 30 ago 2023). https://ecommerce.inn.cl/nch-

iso-iec-27001202386060. Consultado el 17 de agosto de 2026. NORMA

[11] AICPA & CIMA — SOC 2® (System and Organization Controls): informes sobre controles de seguridad, disponibilidad, integridad, confidencialidad y privacidad. https://www.aicpa-cima.com/topic/audit-assurance/audit-and-assurance-greater-than-soc-2. Consultado el 17 de agosto de 2026. NORMA

[12] OWASP Top 10 for LLM Applications 2025 (LLM01-LLM10; edición con traducción al español). https://genai.owasp.org/llm-top-10/. Consultado el 17 de agosto de 2026. NORMA

[13] OWASP — Top 10 2025 de riesgos y mitigaciones para LLMs y aplicaciones de IA generativa (traducción oficial al español, PDF). https://genai.owasp.org/resource/top-10-2025-de-riesgos-y-mitigaciones-para-llms-y-aplicaciones-de-ia-generativa/. Consultado el 17 de agosto de 2026. NORMA

[14] OWASP GenAI LLM Top 10 — 2026 (edición vigente, publicada el 3-4 de agosto de 2026). https://genai.owasp.org/resource/owasp-genai-llm-top-10-2026/. Consultado el 17 de agosto de 2026. NORMA

[15] OWASP Top 10 for Agentic Applications for 2026 (ASI01-ASI10). https://genai.owasp.org/resource/owasp-top-10-for-agentic-applications-for-2026/. Consultado el 17 de agosto de 2026. NORMA

[16] MITRE ATLAS™ — Adversarial Threat Landscape for Artificial-Intelligence Systems. https://atlas.mitre.org/. Consultado el 17 de agosto de 2026. NORMA

[17] Magesh, Surani, Dahl, Suzgun, Manning, Ho — «Hallucination-Free? Assessing the Reliability of Leading AI Legal Research Tools», arXiv 2405.20362 (2024); Journal of Empirical Legal Studies (2025). https://arxiv.org/abs/2405.20362. Consultado el 17 de agosto de 2026. PAPER

[18] OpenAI — Your data (uso de datos de la API: no entrenamiento por defecto desde el 1 mar 2023; registros de monitoreo de abuso hasta 30 días; Zero Data Retention). https://developers.openai.com/api/docs/guides/your-data. Consultado el 17 de agosto de 2026. DOC-FABRICANTE

[19] Anthropic — Commercial Terms of Service («Anthropic may not train models on Customer Content from Services»). https://www.anthropic.com/legal/commercial-terms. Consultado el 17 de agosto de 2026. DOC-FABRICANTE

[20] CMF (Chile) — Recopilación Actualizada de Normas, Capítulo 20-7, Externalización de servicios (bancos). https://www.cmfchile.cl/portal/normativa/624/articles-28982_doc_pdf.pdf. Consultado el 17 de agosto de 2026. NORMA

[21] Innova y Cree — Benchmark anti-alucinación: tasas medidas y publicadas con su metodología (rúbrica Stanford RegLab, IC 95 % Wilson, McNemar pareado, pruebas hostiles). https://innovaycree.com/empresas/benchmark/. Consultado el 17 de agosto de 2026. PROPIO

[22] Ley Federal de Protección de Datos Personales en Posesión de los Particulares (México) — texto vigente, Cámara de Diputados (última reforma D.O.F. 14 nov 2025). https://www.diputados.gob.mx/LeyesBiblio/pdf/LFPDPPP.pdf. Consultado el 18 de agosto de 2026. LEY

[23] Ley 1581 de 2012 (Colombia) — Régimen general de protección de datos personales. http://www.secretariasenado.gov.co/senado/basedoc/ley_1581_2012.html. Consultado el 18 de agosto de 2026. LEY

[24] Ley 25.326 (Argentina) — Protección de los Datos Personales (texto actualizado, InfoLEG). https://servicios.infoleg.gob.ar/infolegInternet/anexos/60000-64999/64790/texact.htm. Consultado el 18 de agosto de 2026. LEY

[25] Reglamento (UE) 2016/679 — Reglamento General de Protección de Datos. https://eur-lex.europa.eu/legal-content/ES/TXT/?uri=CELEX%3A32016R0679. Consultado el 18 de agosto de 2026. LEY

[26] Directiva (UE) 2022/2555 (NIS2) — medidas para un elevado nivel común de ciberseguridad. https://eur-lex.europa.eu/legal-content/ES/TXT/?uri=CELEX%3A32022L2555. Consultado el 18 de agosto de 2026. LEY

[27] Reglamento (UE) 2024/1689 — Reglamento de Inteligencia Artificial. https://eur-lex.europa.eu/legal-content/ES/TXT/?uri=CELEX%3A32024R1689. Consultado el 18 de agosto de 2026. LEY

Citar este artículo

APA

```
Pinto, D. (2026, 17 de agosto). Cómo evaluar a un proveedor de IA: 40 preguntas y evidencia (v1.0). Blog técnico de Innova y Cree.
```

```
https://innovaycree.com/blog/seguridad-ia/como-evaluar-proveedor-ia.html
```

BIBTEX

```
@misc{pinto2026como,
  author       = {Pinto, David},
  title        = {Cómo evaluar a un proveedor de IA: 40 preguntas y evidencia},
  howpublished = {Blog técnico de Innova y Cree},
  year         = {2026},
  month        = {ago},
  note         = {v1.0, revisado 2026-08-17},
  url          = {https://innovaycree.com/blog/seguridad-ia/como-evaluar-
proveedor-ia.html}
}
```

v1.0 · 17 ago 2026

▶ HISTORIAL DE VERSIONES

Descargables de este artículo

- Checklist de evaluación de proveedores de IA (v1.0, XLSX) CC BY-ND 4.0
- Checklist de evaluación de proveedores de IA (v1.0, CSV) CC BY-ND 4.0
- Checklist de evaluación de proveedores de IA (v1.0, Markdown) CC BY-ND 4.0

SOBRE EL AUTOR

David Pinto · Fundador y arquitecto de la plataforma

Diseña la arquitectura de IA verificable de Innova y Cree: RAG con cita verificada contra la fuente oficial, benchmark anti-alucinación público y plataformas en producción para sectores regulados de habla hispana.

LinkedIn · Todos sus artículos