AI puede traducir. Pero, ¿quién traduce el riesgo?

Heather M. Grizzle · August 17, 2026

La IA puede traducir, pero ¿quién traduce el riesgo? Explore la brecha de responsabilidad en la IA, la gobernanza responsable de la IA, las normas de contratación y la supervisión de la IA en lengua de signos.

Contenidos

Toda implementación de IA traslada el riesgo a alguien. En la mayoría de los casos, recae en la persona con menor capacidad para detectar un error y menor autoridad para cuestionarlo.

Heather M. Grizzle Novara Consulting Group 17 de agosto de 2026

Earlier this month I wrote about Google DeepMind's sign language translation research and argued that a demonstration is not the same thing as evidence. The response told me something useful. Almost nobody disagreed that the distinction matters. Many people asked the obvious follow-up question instead: if a demonstration is not enough, who is supposed to decide when a system is good enough to deploy, and who answers for it when it is not?

Esa pregunta no es específica de la lengua de señas. Es el problema central sin resolver en la gobernanza de la IA en este momento, y tiene un nombre que vale la pena usar sin rodeos. Llamémoslo la brecha de rendición de cuentas.

¿Qué es la brecha de rendición de cuentas en la IA?

The AI accountability gap is the distance between what a system is capable of doing and what any identifiable party is answerable for when it does that thing badly. It opens whenever an organization deploys an AI system faster than it builds the means to verify the system's output, hear from the people the output affects, and correct or stop the deployment when the output is wrong.

La brecha no es principalmente un problema técnico. Los modelos fallan de maneras predecibles, y los ingenieros suelen ser francos sobre esos modos de fallo. La brecha es un problema institucional. Las organizaciones están adquiriendo sistemas cuya producción no pueden evaluar, implementándolos en entornos donde los errores tienen consecuencias, y sin conservar ningún mecanismo práctico para averiguar si el sistema está funcionando.

Capacidad, fiabilidad y rendición de cuentas son tres tipos distintos de evidencia

La mayoría de las conversaciones de adquisición tratan esto como una sola cosa. No lo es, y separarlas es el movimiento más útil que puede hacer un comprador institucional.

Capability evidence shows that a system can perform a task under conditions the vendor selected. Demonstrations, benchmark scores, and pilot footage are capability evidence. They answer the question "is this possible."

Reliability evidence shows how the system performs across the range of people, settings, and conditions the buyer will actually encounter, including the ones the vendor did not choose. Disaggregated performance data, independent testing, and error rates from live deployment are reliability evidence. They answer the question "how often is this wrong, and for whom."

Accountability evidence shows what happens when the system fails. It names who monitors performance, who receives complaints, who has authority to suspend the deployment, what the affected person is owed, and how any of that is enforced. It answers the question "who is responsible."

Vendors compete on capability evidence because it is the cheapest to produce and the most persuasive to watch. Reliability evidence is expensive and unflattering. Accountability evidence is not really the vendor's to produce at all, because it describes the buyer's own governance, and most buyers have not built it.

Una organización que acepta la evidencia de capacidad como si resolviera las otras dos cuestiones no ha tomado una decisión de compra. Ha hecho una suposición y la ha pagado.

Por qué las leyes de divulgación no cierran la brecha

La respuesta regulatoria más visible al riesgo de la IA en este momento es la transparencia. Según el artículo 50 de la Ley de IA de la UE (Unión Europea, Ley de Inteligencia Artificial), las obligaciones que entraron en vigor el 2 de agosto de 2026 exigen que los proveedores marquen el contenido sintético en formato legible por máquina y que los implementadores divulguen las deepfakes y ciertos textos generados por IA sobre asuntos de interés público, con sanciones que llegan hasta 15 millones de euros o el tres por ciento de la facturación anual mundial total. La Comisión Europea concedió un período de gracia hasta el 2 de diciembre de 2026 para los sistemas generativos ya comercializados.

Este es un trabajo significativo y lo apoyo. También vale la pena precisar qué hace exactamente. El artículo 50 es un régimen de procedencia. Le dice a una persona que una máquina intervino. No le dice si la máquina acertó.

For a great deal of synthetic media, provenance is the whole question. If the concern is a fabricated video of a public figure, knowing the video is synthetic resolves it. But for AI systems that mediate communication, provide information, or influence a deci-

sion about a person, provenance is the easy half. A label reading "this translation was generated by AI" leaves the recipient exactly where they started, which is unable to tell whether the content is accurate.

The liability half of the European approach has moved in the opposite direction. The Commission's proposed AI Liability Directive, which would have eased the burden of proof for people harmed by AI systems, was formally withdrawn in October 2025. Transparency advanced. Redress did not.

El patrón se repite en Estados Unidos a nivel estatal. Colorado aprobó la primera ley estatal integral sobre IA en 2024, y luego la derogó y sustituyó por la SB 26-189, firmada el 14 de mayo de 2026, cuyas obligaciones sustantivas para desarrolladores e implementadores se aplican a partir del 1 de enero de 2027. La sustitución eliminó los programas obligatorios de gestión de riesgos, las evaluaciones de impacto algorítmico y el deber independiente de diligencia razonable frente a la discriminación algorítmica. Lo que sobrevivió fue el aviso previo al uso, una divulgación en lenguaje sencillo dentro de los treinta días posteriores a una decisión adversa, la conservación de registros durante tres años y el derecho a solicitar una revisión humana significativa.

No estoy argumentando que esas obligaciones supervivientes carezcan de valor. El aviso y la conservación de registros son reales. Pero el aviso te dice que se usó un sistema. Los registros te dicen qué hizo. Ninguno establece que la salida fue precisa, y ninguno asigna responsabilidad por las consecuencias si no lo fue. La dirección del recorrido regulatorio en los últimos dos años ha sido hacia informar a las personas de que la IA intervino y alejarse de exigir que alguien responda por lo que produjo.

Asimetría de verificación: el problema estructural subyacente

Aquí está la parte que considero poco examinada, y es donde el caso de la lengua de señas deja de ser un ejemplo marginal y se convierte en la ilustración más clara disponible de una condición general.

En la mayoría de las implementaciones de IA, la persona mejor posicionada para detectar un error no tiene autoridad para actuar sobre él, y la parte con autoridad para actuar no tiene los medios para detectarlo. Yo lo llamaría asimetría de verificación, y es el mecanismo que produce la brecha de rendición de cuentas.

Considere lo que ve una institución oyente cuando implementa un sistema automatizado de lengua de señas. Ve que el sistema produjo una salida. El personal observa un avatar haciendo señas o aparecen subtítulos. Desde ese punto de vista, la transacción parece completa.

Now consider what the Deaf person experiences. They receive output they cannot check against the source, delivered by a system they did not select, in a setting where the institution has already concluded that access was provided. If the rendering drops a negation, flattens a conditional, misplaces spatial reference, or omits the non-manual markers that carry grammatical meaning, the person may not know. If they do notice something is wrong, they are in the position of contesting the institution′s own conclusion that communication succeeded.

Eso no es una queja sobre tecnología. Es una distribución de poder. La institución tiene la autoridad para determinar si hubo acceso y carece de la capacidad lingüística para evaluarlo. La persona Deaf tiene la capacidad lingüística y carece de la autoridad. El error, si lo hay, no tiene dónde salir a la luz.

Una vez que se ve la forma que tiene, se encuentra casi en todas partes. Un paciente no puede evaluar una recomendación de triaje asistida por IA. Un solicitante no puede inspeccionar el modelo que lo descartó. Un beneficiario de prestaciones no puede auditar la determinación de elegibilidad. Una persona que recibe instrucciones legales traducidas por máquina en cualquier idioma no puede verificar la traducción sin la fluidez que la traducción existe para suplir. En cada caso, la parte afectada absorbe un riesgo que no puede medir, y la institución que implementa registra una transacción completada.

Lo que hace que la lengua de señas sea un caso particularmente agudo son las consecuencias asociadas al contenido. El lenguaje en entornos institucionales conlleva consentimiento, testimonio, diagnóstico, instrucción y efecto legal. Una traducción errónea no es una experiencia de usuario degradada. Es un formulario de consentimiento defectuoso, un historial médico inexacto o una declaración tergiversada en un procedimiento.

La ley de accesibilidad ya respondió parte de esto, y lo estamos desaprendiendo

Ya existe una respuesta viable en la práctica de accesibilidad, lo que hace más difícil de disculpar su descuido actual.

Under Department of Justice regulation at 28 CFR 35.160, a public entity must give primary consideration to the auxiliary aid or service requested by the person with a disability. The rule for private public accommodations at 28 CFR 36.303 is weaker, requiring consultation while leaving the final choice with the provider, and that difference is itself instructive about where verification authority tends to erode. The National Association of the Deaf's minimum standards for video remote interpreting in medical settings state the underlying principle directly: a deaf or hard of hearing individual knows best which auxiliary aid or service will achieve effective communication. Those same standards go further and require the video interpreter to tell the provider to terminate the remote session and obtain an on-site interpreter when the interpreter determines the remote arrangement is not producing effective communication.

Léalo como un diseño de gobernanza en lugar de una norma de accesibilidad, y resulta inusualmente bien construido. Sitúa el juicio de verificación en la persona que realmente puede emitirlo. Otorga a un profesional presente la autoridad para detener una implementación en pleno uso. Trata la efectividad como algo que se determina en contexto, en lugar de asumirse por la mera presencia del equipo.

Automated systems tend to strip both features out. There is no qualified interpreter positioned to call a halt, because the interpreter is what the system replaced. And the affected person's judgment about whether communication worked is quietly demoted from a legal standard to customer feedback.

La Federación Mundial de Sordos (World Federation of the Deaf) y la Asociación Mundial de Intérpretes de Lengua de Señas (World Association of Sign Language Interpreters) establecieron un límite sobre los avatares de señas ya en 2018. Su postura fue que los avatares son inapropiados para información en vivo, compleja o significativa, y que el contenido estático preregrabado puede ser aceptable cuando personas Deaf asesoraron sobre el material y no se requiere interacción en vivo. Esa declaración tiene ocho años. La tecnología ha avanzado considerablemente desde entonces. El límite de implementación que trazó no ha sido reemplazado por nada más riguroso, y sigue siendo la línea más clara disponible.

Por qué la demostración del proveedor no puede ser el registro de gobernanza

La demostración está diseñada para ser persuasiva. Esa es su función, y no hay nada indebido en ello. El problema es que, en ausencia de una evaluación independiente, la demostración se convierte por defecto en la base probatoria de una decisión de adquisición, porque es el único documento presente.

Una demostración establece la posibilidad bajo condiciones seleccionadas. No puede establecer el desempeño en toda la población a la que sirve el comprador, el comportamiento fuera de condiciones controladas, la seguridad en entornos de altas consecuencias, ni la rendición de cuentas cuando ocurren errores. Esas son propiedades de una implementación, no propiedades de un modelo.

La política federal ya reconoce esto en principio. La guía de la OMB (Oficina de Administración y Presupuesto, Office of Management and Budget) emitida en abril de 2025 define la IA de alto impacto como los sistemas cuya salida sirve como base principal para decisiones o acciones con un efecto legal, material, vinculante o significativo sobre derechos o seguridad, y el memorando enumera ámbitos que incluyen derechos civiles, acceso a la educación, vivienda, empleo, beneficios y servicios gubernamentales, y salud. Para esos usos exige pruebas previas a la implementación, evaluaciones de impacto de la IA antes de la implementación y periódicamente después, supervisión humana adecuada, y revisión humana oportuna con posibilidad de apelación para las personas afectadas. Las agencias que no puedan cumplir esos mínimos deben interrumpir el uso de forma segura. El memorando complementario sobre adquisiciones exige cláusulas contractuales que otorguen a la agencia la capacidad de monitorear y evaluar regularmente el desempeño, los riesgos y la efectividad del sistema.

Sea cual sea la opinión que se tenga sobre los detalles, la estructura de esa guía es correcta. Trata la rendición de cuentas como algo especificado en el contrato, probado antes del lanzamiento, monitoreado después del lanzamiento y exigible mediante la suspensión. Eso es una capa de verificación. La mayoría de las instituciones que compran IA fuera del sistema federal de adquisiciones no tienen nada equivalente.

Añadiría una advertencia sobre asumir que la capacidad llegará según lo previsto y cerrará la brecha por sí sola. En abril de 2026, el Departamento de Justicia retrasó los plazos de cumplimiento de la norma de accesibilidad web del Título II de la ADA (Ley para Estadounidenses con Discapacidades, Americans with Disabilities Act) en apro-

ximadamente un año para cada categoría de entidad, moviendo a las entidades públicas grandes al 26 de abril de 2027 y a las más pequeñas al 26 de abril de 2028. Entre las razones que dio el Departamento estaba que había sobreestimado las capacidades, tanto de personal como de tecnología, de las entidades cubiertas, y afirmó claramente que la tecnología avanzada, como la IA generativa, todavía no automatiza de forma fiable la remediación de contenido inaccesible a gran escala. Eso es un regulador federal dejando constancia de que las afirmaciones sobre la capacidad de la IA superaron su cumplimiento real en un contexto de accesibilidad. Vale la pena detenerse en ello antes de tratar cualquier calendario de un proveedor como un plan de gobernanza.

Lo que las instituciones deberían exigir antes de la implementación

Las preguntas que siguen no son exóticas. Son las preguntas que un comprador competente haría sobre cualquier sistema que conlleve consecuencias, y son respondibles.

¿Quién evaluó este sistema, y era independiente del proveedor y del comprador?

¿Qué poblaciones, dialectos, variantes regionales y entornos se incluyeron en las pruebas, y cuáles fueron los resultados para cada uno en lugar de en conjunto?

¿Con qué datos se entrenó el sistema, bajo qué licencia, y con qué consentimiento de las personas cuyo lenguaje capturó?

¿Cuál es la tasa de error medida en condiciones similares a las nuestras, y qué ocurre con esa tasa cuando las condiciones empeoran?

¿Quién en nuestra organización tiene la autoridad para suspender esta implementación, y qué activa esa autoridad?

¿Cómo reporta una persona afectada que el resultado fue incorrecto, quién recibe ese reporte, y qué cambia como resultado?

¿Qué tenemos derecho contractual a saber sobre el rendimiento después de la puesta en marcha, y cuál es nuestro recurso si el rendimiento disminuye?

Una organización que no puede responder estas preguntas no ha evaluado un sistema. Lo ha aceptado. Novara Consulting Group creó el Sign Language Access Trust Index para hacer que estas preguntas sean respondibles en una forma que las adquisiciones puedan realmente utilizar, con estándares de evidencia definidos, puntuación a nivel

de indicador, y una distinción entre lo que un proveedor afirma y lo que una parte independiente puede confirmar. El instrumento es específico para la IA de lenguaje de señas. La disciplina subyacente no lo es.

La siguiente etapa de la IA es la verificación

La conversación pública sobre la IA se ha organizado en torno a la capacidad durante tres años. Puede escribir, puede razonar, puede ver, puede actuar, puede interpretar la comunicación humana. Esas preguntas se están respondiendo rápidamente, y a menudo de forma afirmativa.

La pregunta que determina si algo de esto es confiable es diferente. Es si las instituciones que implementan estos sistemas pueden saber cuándo los sistemas están equivocados, y si las personas afectadas por ellos tienen alguna posición para decirlo.

En este momento, en la mayoría de las implementaciones, la respuesta a ambas es no. El riesgo no ha sido traducido. Ha sido transferido a quien tiene menos capacidad de verlo.

Closing that gap is not a matter of waiting for better models. It requires evaluation independent of the vendor, evidence standards that survive contact with procurement, authority to stop a deployment located with someone in the room, and a channel through which the affected person's judgment counts for something. None of that is technically hard. It is institutionally inconvenient, which is a different problem and a more tractable one.

Preguntas frecuentes

¿Qué es la brecha de responsabilidad de la IA? La brecha de responsabilidad de la IA es la distancia entre lo que un sistema de IA es capaz de hacer y de qué es responsable cualquier parte identificable cuando lo hace mal. Aparece cuando las organizaciones implementan IA más rápido de lo que construyen la capacidad de verificar su resultado, recibir quejas de las personas afectadas, y suspender la implementación cuando falla.

¿Es la transparencia de la IA lo mismo que la responsabilidad de la IA? No. Las reglas de transparencia, incluido el Artículo 50 de la Ley de IA de la UE, generalmente establecen la procedencia, es decir, informan a las personas de que el conte-

nido fue generado o manipulado por IA. La responsabilidad establece quién es responsable de la exactitud de ese contenido y qué se le debe a la persona afectada cuando está equivocado. Un sistema puede estar completamente etiquetado y aun así no ser responsable.

¿Quién es legalmente responsable cuando un sistema de IA causa daño? En la mayoría de las jurisdicciones esto sigue sin resolverse. La Comisión Europea retiró su propuesta de Directiva de Responsabilidad por IA en octubre de 2025, y varios marcos estatales de Estados Unidos se han reducido hacia obligaciones de notificación y mantenimiento de registros en lugar de deberes sustantivos de cuidado. En la práctica, la responsabilidad se determina por los términos contractuales, la legislación específica del sector como los estatutos de discapacidad y derechos civiles, y la doctrina general de responsabilidad, en lugar de por reglas específicas de IA.

Why is sign language AI a useful test case for AI governance? Because it makes verification asymmetry visible. The institution deploying the system usually cannot evaluate the output, and the Deaf person who can evaluate it usually has no authority to challenge the institution's conclusion that access was provided. The same structure operates in medical, hiring, and benefits AI, but is harder to observe.

¿Qué debería exigir una organización antes de implementar un sistema de IA que afecta a las personas? Evaluación independiente en lugar de demostración del proveedor, datos de rendimiento desglosados por población y entorno, procedencia divulgada de los datos de entrenamiento, una persona designada con autoridad para suspender la implementación, un canal de quejas funcional para las personas afectadas, y derechos contractuales para monitorear el rendimiento después del lanzamiento.

What is the SLAT Index? The Sign Language Access Trust Index is Novara Consulting Group's evaluation instrument for sign language AI systems. It sets defined evidence standards across governance, linguistic, transparency, accessibility, and deployment domains, scores systems at the indicator level, and distinguishes vendor assertion from independently confirmable fact so that procurement decisions rest on evidence rather than demonstration.

Heather M. Grizzle es la fundadora de Novara Consulting Group, que asesora a instituciones sobre la gobernanza y adquisición de tecnología de accesibilidad de IA. Es la autora del SLAT Index Evaluation Standard.

Related reading: The Demo Is Not the Evidence: What Google DeepMind's Sign Language AI Must Prove Next

Fuentes

- Comisión Europea, Datos rápidos: Reglas de transparencia para sistemas de IA (Artículo 50, aplicable a partir del 2 de agosto de 2026). https://digital-strategy.ec.europa.eu/en/factpages/quick-facts-transparency-rules-ai-systems
- Calendario legislativo del Parlamento Europeo, Directiva de Responsabilidad por IA (retiro, octubre de 2025). https://www.europarl.europa.eu/legislative-train/theme-a-europe-fit-for-the-digital-age/file-ai-liability-directive
- Departamento de Justicia de EE. UU., Extensión de las fechas de cumplimiento para la no discriminación por motivos de discapacidad: Accesibilidad de la información y servicios web de las entidades gubernamentales estatales y locales, 20 de abril de 2026. https://www.federalregister.gov/documents/2026/04/20/2026-07663/extension-of-compliance-dates-for-nondiscrimination-on-the-basis-of-disability-accessibility-of-web
- Oficina de Administración y Presupuesto, M-25-21, Acelerando el Uso Federal de la IA a través de la Innovación, la Gobernanza y la Confianza Pública (abril de 2025). https://www.whitehouse.gov/wp-content/uploads/2025/02/M-25-21-Accelerating-Federal-Use-of-AI-through-Innovation-Governance-and-Public-Trust.pdf
- Oficina de Administración y Presupuesto, M-25-22, Impulsando la Adquisición Eficiente de Inteligencia Artificial en el Gobierno (abril de 2025). https://www.whitehouse.gov/wp-content/uploads/2025/02/M-25-22-Driving-Efficient-Acquisition-of-Artificial-Intelligence-in-Government.pdf
- Colorado SB 26-189, que deroga y vuelve a promulgar el marco de la SB 24-205, firmada el 14 de mayo de 2026, con obligaciones sustantivas aplicables a partir del 1 de enero de 2027. https://leg.colorado.gov/bills/sb26-189
- 28 CFR 35.160, Comunicaciones (Título II, consideración primordial). https://www.ecfr.gov/current/title-28/chapter-I/part-35/subpart-E/section-35.160

- 28 CFR 36.303, Ayudas y servicios auxiliares (Título III, estándar de consulta). https://www.ecfr.gov/current/title-28/chapter-I/part-36/subpart-C/section-36.303
- Asociación Nacional de Deaf, Estándares Mínimos para los Servicios de Interpretación Remota por Video en Entornos Médicos. https://nad.org/minimum-standards-for-video-remote-interpreting-services-in-medical-settings/
- Federación Mundial de Deaf y Asociación Mundial de Intérpretes de Lengua de Señas, Declaración sobre el Uso de Avatares de Señas, 14 de marzo de 2018, actualizada el 14 de abril de 2018. https://wasli.org/wp-content/uploads/2023/07/WFD-and-WASLI-Statement-on-Avatar-FINAL-14032018-Updated-14042018-1.pdf

Cómo citar

Grizzle, H. M. (2026, August 17). AI puede traducir. Pero, ¿quién traduce el riesgo?. Novara Consulting Group. https://www.novaracg.com/es/2026/08/17/ai-can-translate-but-who-translates-the-risk/