This document was translated from English by machine and may contain mistakes. The English original is at https://www.novaracg.com/2026/08/17/ai-can-translate-but-who-translates-the-risk/.

L'IA può tradurre. Ma chi traduce il rischio?

Heather M. Grizzle · August 17, 2026

AI can translate, but who translates the risk? The AI accountability gap, procurement standards, and why verification falls on the person least able to check.

Contenuti

1. Cos'è il divario di responsabilità nell'IA?
2. Capacità, affidabilità e responsabilità sono tre tipi diversi di prova
3. Perché le leggi sulla divulgazione non colmano il divario
4. Asimmetria di verifica: il problema strutturale sottostante
5. La legge sull'accessibilità ha già risposto in parte, e lo stiamo disimparando
6. Perché la dimostrazione del fornitore non può essere il documento di governance
7. Cosa dovrebbero richiedere le istituzioni prima della messa in servizio
8. La prossima fase dell'IA è la verifica
9. Domande frequenti
10. Fonti

Ogni distribuzione dell'IA trasferisce il rischio su qualcuno. Nella maggior parte dei casi ricade sulla persona con la minore capacità di rilevare un errore e la minore autorità per contestarlo.
Heather M. Grizzle Novara Consulting Group, 17 agosto 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?

Questa domanda non riguarda solo la lingua dei segni. È il problema centrale e irrisolto della governance dell'IA in questo momento, e ha un nome che vale la pena usare apertamente. Chiamiamolo il divario di responsabilità.

Cos'è il divario di responsabilità nell'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.

Il divario non è principalmente un problema tecnico. I modelli falliscono in modi prevedibili, e gli ingegneri sono di solito sinceri su queste modalità di fallimento. Il divario è un problema istituzionale. Le organizzazioni stanno acquisendo sistemi i cui output non sono in grado di valutare, distribuendoli in contesti in cui gli errori comportano conseguenze, e non mantenendo alcun meccanismo pratico per scoprire se il sistema funziona.

Capacità, affidabilità e responsabilità sono tre tipi diversi di prova

La maggior parte delle discussioni sugli appalti tratta questi elementi come un'unica cosa. Non lo sono, e separarli è la mossa più utile che un acquirente istituzionale possa compiere.

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.

Un'organizzazione che accetta le prove di capacità come se risolvessero anche le altre due questioni non ha preso una decisione di acquisto. Ha fatto un'ipotesi e l'ha pagata.

Perché le leggi sulla divulgazione non colmano il divario

La risposta normativa più visibile al rischio dell'IA in questo momento è la trasparenza. Ai sensi dell'articolo 50 dell'EU AI Act (regolamento sull'intelligenza artificiale dell'Unione Europea), gli obblighi divenuti applicabili il 2 agosto 2026 richiedono ai fornitori di contrassegnare i contenuti sintetici in forma leggibile da macchina e richiedono ai responsabili della distribuzione di segnalare i deepfake e alcuni testi generati dall'IA su questioni di interesse pubblico, con sanzioni fino a 15 milioni di euro o il tre percento del fatturato mondiale totale annuo. La Commissione Europea ha concesso un periodo di grazia fino al 2 dicembre 2026 per i sistemi generativi già immessi sul mercato.

Si tratta di un lavoro significativo e lo sostengo. Vale anche la pena essere precisi su cosa comporti. L'articolo 50 è un regime di provenienza. Dice a una persona che una macchina è stata coinvolta. Non le dice se la macchina aveva ragione.

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 decision 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.

Lo schema si ripete negli Stati Uniti a livello statale. Il Colorado ha approvato la prima legge statale completa sull'IA nel 2024, per poi abrogarla e sostituirla con l'SB 26-189, firmata il 14 maggio 2026, i cui obblighi sostanziali per sviluppatori e responsabili della distribuzione si applicano a partire dal 1° gennaio 2027. La sostituzione ha eliminato i programmi obbligatori di gestione del rischio, le valutazioni d'impatto algoritmico e il dovere autonomo di ragionevole diligenza contro la discriminazione algoritmica. Ciò che è sopravvissuto sono il preavviso prima dell'uso, una divulgazione in linguaggio semplice entro trenta giorni da una decisione sfavorevole, la conservazione dei documenti per tre anni e un diritto a richiedere una revisione umana significativa.

Non sto sostenendo che questi obblighi sopravvissuti siano privi di valore. Il preavviso e la conservazione dei documenti sono elementi reali. Ma il preavviso ti dice che un sistema è stato utilizzato. I documenti ti dicono cosa ha fatto. Nessuno dei due stabilisce che l'output fosse accurato, e nessuno dei due assegna la responsabilità per le conseguenze se non lo era. La direzione del cammino normativo negli ultimi due anni è stata quella di dire alle persone che l'IA era coinvolta, allontanandosi dal far rispondere qualcuno di ciò che ha prodotto.

Asimmetria di verifica: il problema strutturale sottostante

Ecco la parte che ritengo sia poco esaminata, ed è il punto in cui il caso della lingua dei segni smette di essere un esempio di nicchia e diventa l'illustrazione più chiara disponibile di una condizione generale.

Nella maggior parte delle distribuzioni di IA, la persona meglio posizionata per rilevare un errore non ha l'autorità di agire su di esso, e la parte con l'autorità di agire non ha i mezzi per rilevarlo. La chiamerei asimmetria di verifica, ed è il meccanismo che produce il divario di responsabilità.

Si consideri cosa vede un'istituzione udente quando distribuisce un sistema automatizzato di lingua dei segni. Vede che il sistema ha prodotto un output. Il personale osserva un avatar che firma o sottotitoli che appaiono. Da questo punto di vista la transazione sembra 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.

Non si tratta di una lamentela sulla tecnologia. È una distribuzione di potere. L'istituzione detiene l'autorità di determinare se l'accesso sia avvenuto e manca della capacità linguistica per valutarlo. La persona Deaf detiene la capacità linguistica e manca dell'autorità. L'errore, se esiste, non ha dove emergere.

Una volta che se ne vede la forma, la si trova quasi ovunque. Un paziente non può valutare una raccomandazione di triage assistita dall'IA. Un candidato non può ispezionare il modello che lo ha scartato. Un richiedente di benefici non può verificare la determinazione di idoneità. Una persona che riceve istruzioni legali tradotte automaticamente in qualsiasi lingua non può verificare la traduzione senza la padronanza linguistica che la traduzione esiste per fornire. In ogni caso la parte colpita assorbe un rischio che non può misurare, e l'istituzione che effettua la distribuzione registra una transazione completata.

Ciò che rende la lingua dei segni un caso particolarmente netto è la posta in gioco legata al contenuto. Il linguaggio nei contesti istituzionali veicola consenso, testimonianza, diagnosi, istruzione ed effetto legale. Una traduzione errata non è un'esperienza utente degradata. È un modulo di consenso difettoso, un'anamnesi medica imprecisa, o una dichiarazione travisata in un procedimento.

La legge sull'accessibilità ha già risposto in parte, e lo stiamo disimparando

Esiste già una risposta praticabile nella pratica dell'accessibilità, il che rende più difficile scusare la sua attuale trascuratezza.

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 differen-

ce 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.

Letto come un progetto di governance piuttosto che come una norma sull'accessibilità, è insolitamente ben costruito. Colloca il giudizio di verifica presso la persona che può effettivamente formularlo. Dà a un professionista presente nella stanza l'autorità di interrompere una distribuzione a metà del suo utilizzo. Tratta l'efficacia come qualcosa da determinare nel contesto piuttosto che presumerla dalla presenza di apparecchiature.

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 World Federation of the Deaf e la World Association of Sign Language Interpreters hanno fissato un limite agli avatar firmanti già nel 2018. La loro posizione era che gli avatar sono inappropriati per informazioni dal vivo, complesse o significative, e che i contenuti statici preregistrati possono essere accettabili laddove le persone Deaf abbiano fornito consulenza sul materiale e non sia richiesta alcuna interazione dal vivo. Questa dichiarazione risale a otto anni fa. La tecnologia è avanzata considerevolmente da allora. Il limite di distribuzione che tracciava non è stato sostituito da nulla di più rigoroso, ed è ancora la linea più chiara disponibile.

Perché la dimostrazione del fornitore non può essere il documento di governance

La dimostrazione è progettata per essere persuasiva. Questa è la sua funzione, e non c'è nulla di improprio in ciò. Il problema è che, in assenza di una valutazione indipendente, la dimostrazione diventa per impostazione predefinita la base probatoria per una decisione di appalto, perché è l'unico documento presente nella stanza.

Una dimostrazione stabilisce la possibilità in condizioni selezionate. Non può stabilire le prestazioni nell'intera popolazione servita dall'acquirente, il comportamento al di fuori di condizioni controllate, la sicurezza in contesti ad alte conseguenze, o la responsabilità quando si verificano errori. Queste sono proprietà di una distribuzione, non proprietà di un modello.

La politica federale lo riconosce già in linea di principio. Le linee guida dell'OMB (Office of Management and Budget) emesse nell'aprile 2025 definiscono l'IA ad alto impatto come sistemi il cui output costituisce la base principale per decisioni o azioni con un effetto legale, materiale, vincolante o significativo su diritti o sicurezza, con il promemoria che enumera ambiti tra cui diritti civili, accesso all'istruzione, alloggio, occupazione, benefici e servizi governativi, e salute. Per questi usi richiede test pre-distribuzione, valutazioni d'impatto dell'IA prima della distribuzione e periodicamente in seguito, una supervisione umana adeguata, e una tempestiva revisione umana con possibilità di ricorso per le persone colpite. Le agenzie che non riescono a soddisfare questi requisiti minimi devono interrompere in sicurezza l'uso. Il promemoria complementare sugli appalti richiede clausole contrattuali che diano all'agenzia la capacità di monitorare e valutare regolarmente le prestazioni, i rischi e l'efficacia del sistema.

Qualunque cosa si pensi dei dettagli, la struttura di quelle linee guida è corretta. Tratta la responsabilità come qualcosa di specificato nel contratto, testato prima del lancio, monitorato dopo il lancio, e applicabile tramite sospensione. Si tratta di uno strato di verifica. La maggior parte delle istituzioni che acquistano IA al di fuori del sistema di acquisizione federale non ha nulla di equivalente.

Aggiungerei una cautela sul presupposto che la capacità arrivi secondo i tempi previsti e colmi il divario da sola. Nell'aprile 2026 il Dipartimento di Giustizia ha posticipato le scadenze di conformità per la norma sull'accessibilità web del Titolo II dell'ADA (Americans with Disabilities Act) di circa un anno per ogni categoria di ente, spostando gli enti pubblici grandi al 26 aprile 2027 e quelli più piccoli al 26 aprile 2028. Tra le ragioni addotte dal Dipartimento vi era il fatto di aver sovrastimato le capacità, in termini di personale e di tecnologia, degli enti soggetti all'obbligo, e ha affermato chiaramente che tecnologie avanzate come l'IA generativa non automatizzano ancora in modo affidabile la rimozione di contenuti inaccessibili su larga scala. Si tratta di un regolatore federale che registra come le affermazioni sulle capacità dell'IA abbiano supe-



rato in velocità i risultati effettivi in un contesto di accessibilità. Vale la pena rifletterci prima di trattare qualsiasi tempistica di un fornitore come un piano di governance.

Cosa dovrebbero richiedere le istituzioni prima della messa in servizio

Le domande seguenti non sono esotiche. Sono le domande che un acquirente competente porrebbe su qualsiasi sistema che comporti conseguenze, e sono domande a cui si può rispondere.

Chi ha valutato questo sistema, ed era indipendente dal fornitore e dall'acquirente?

Quali popolazioni, dialetti, varianti regionali e contesti sono stati inclusi nei test, e quali sono stati i risultati per ciascuno, anziché in aggregato?

Su quali dati è stato addestrato il sistema, con quale licenza, e con quale consenso da parte delle persone il cui linguaggio è stato raccolto?

Qual è il tasso di errore misurato in condizioni simili alle nostre, e cosa succede a quel tasso quando le condizioni peggiorano?

Chi nella nostra organizzazione ha l'autorità di sospendere questa implementazione, e cosa attiva tale autorità?

Come fa una persona interessata a segnalare che l'output era errato, chi riceve tale segnalazione, e cosa cambia di conseguenza?

A cosa abbiamo diritto contrattualmente di sapere sulle prestazioni dopo l'avvio operativo, e qual è il nostro rimedio se le prestazioni calano?

Un'organizzazione che non può rispondere a queste domande non ha valutato un sistema. Lo ha accettato. Novara Consulting Group ha creato lo Sign Language Access Trust Index per rendere queste domande rispondibili in una forma che gli uffici acquisti possano effettivamente utilizzare, con standard di evidenza definiti, punteggi a livello di indicatore e una distinzione tra ciò che un fornitore afferma e ciò che una parte indipendente può confermare. Lo strumento è specifico per l'IA del linguaggio dei segni. La disciplina sottostante non lo è.

La prossima fase dell'IA è la verifica

Il dibattito pubblico sull'IA è stato organizzato attorno alla capacità per tre anni. Sa scrivere, sa ragionare, sa vedere, sa agire, sa interpretare la comunicazione umana. A queste domande si sta rispondendo rapidamente, e spesso in modo affermativo.

La domanda che determina se tutto ciò sia affidabile è diversa. È se le istituzioni che implementano questi sistemi possano accorgersi di quando i sistemi sbagliano, e se le persone interessate abbiano un qualche diritto di farlo presente.

Al momento, nella maggior parte delle implementazioni, la risposta a entrambe è no. Il rischio non è stato tradotto. È stato trasferito a chi è meno in grado di percepirlo.

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.

Domande frequenti

Cos'è il divario di responsabilità dell'IA? Il divario di responsabilità dell'IA è la distanza tra ciò che un sistema di IA è in grado di fare e ciò di cui qualsiasi parte identificabile è responsabile quando lo fa male. Appare quando le organizzazioni implementano l'IA più velocemente di quanto costruiscano la capacità di verificarne l'output, ricevere reclami dalle persone interessate e sospendere l'implementazione quando fallisce.

La trasparenza dell'IA è la stessa cosa della responsabilità dell'IA? No. Le norme sulla trasparenza, incluso l'Articolo 50 dell'EU AI Act, generalmente stabiliscono la provenienza, cioè informano le persone che il contenuto è stato generato o manipolato dall'IA. La responsabilità stabilisce chi risponde dell'accuratezza di quel contenuto e cosa è dovuto alla persona interessata quando è errato. Un sistema può essere completamente etichettato e rimanere comunque privo di responsabilità.

Chi è legalmente responsabile quando un sistema di IA causa un danno? Nella maggior parte delle giurisdizioni questo resta irrisolto. La Commissione Europea ha ritirato la proposta di Direttiva sulla responsabilità civile dell'IA nell'ottobre

2025, e diversi quadri normativi statali statunitensi si sono ristretti verso obblighi di notifica e tenuta dei registri piuttosto che doveri di diligenza sostanziali. Nella pratica, la responsabilità è determinata dai termini contrattuali, dalla legge specifica del settore come gli statuti sulla disabilità e i diritti civili, e dalla dottrina generale della responsabilità civile piuttosto che da norme specifiche per l'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.

Cosa dovrebbe richiedere un'organizzazione prima di implementare un sistema di IA che riguarda le persone? Una valutazione indipendente anziché una dimostrazione del fornitore, dati sulle prestazioni disaggregati per popolazione e contesto, provenienza divulgata dei dati di addestramento, una persona nominata con l'autorità di sospendere l'implementazione, un canale funzionante di reclamo per le persone interessate, e diritti contrattuali di monitorare le prestazioni dopo il lancio.

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 è la fondatrice di Novara Consulting Group, che consiglia le istituzioni sulla governance e l'acquisto di tecnologia di accessibilità basata sull'IA. È l'autrice dello SLAT Index Evaluation Standard.

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

---

Fonti

- Commissione Europea, Quick Facts: norme di trasparenza per i sistemi di IA (Articolo 50, applicabile dal 2 agosto 2026). https://digital-strategy.ec.europa.eu/en/factpages/quick-facts-transparency-rules-ai-systems
- Calendario Legislativo del Parlamento Europeo, Direttiva sulla responsabilità civile dell'IA (ritiro, ottobre 2025). https://www.europarl.europa.eu/legislative-train/theme-a-europe-fit-for-the-digital-age/file-ai-liability-directive
- Dipartimento di Giustizia degli Stati Uniti, Proroga delle scadenze di conformità per la non discriminazione sulla base della disabilità: accessibilità delle informazioni e dei servizi web degli enti governativi statali e locali, 20 aprile 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
- Ufficio per la Gestione e il Bilancio, M-25-21, Accelerare l'uso federale dell'IA attraverso innovazione, governance e fiducia pubblica (aprile 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
- Ufficio per la Gestione e il Bilancio, M-25-22, Guidare l'acquisizione efficiente dell'intelligenza artificiale nel governo (aprile 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, che abroga e riformula il quadro normativo SB 24-205, firmato il 14 maggio 2026, con obblighi sostanziali applicabili dal 1° gennaio 2027. https://leg.colorado.gov/bills/sb26-189
- 28 CFR 35.160, Comunicazioni (Titolo II, considerazione primaria). https://www.ecfr.gov/current/title-28/chapter-I/part-35/subpart-E/section-35.160
- 28 CFR 36.303, Ausili e servizi ausiliari (Titolo III, standard di consultazione). https://www.ecfr.gov/current/title-28/chapter-I/part-36/subpart-C/section-36.303
- National Association of the Deaf, Standard minimi per i servizi di interpretariato video a distanza in contesti medici. https://nad.org/minimum-standards-for-video-remote-interpreting-services-in-medical-settings/
- World Federation of the Deaf e World Association of Sign Language Interpreters, Dichiarazione sull'uso di avatar per la lingua dei segni, 14 marzo

2018, aggiornata il 14 aprile 2018. https://wasli.org/wp-content/uploads/2023/07/WFD-and-WASLI-Statement-on-Avatar-FINAL-14032018-Updated-14042018-1.pdf

TIPS

Know something about a sign language AI system that the public should know? Send it to tips@novaracg.com. Every tip is read by me and kept confidential. If you work in the UK, writing to us is not by itself a protected disclosure under the Public Interest Disclosure Act 1998, so before you share employer information, check your options with Protect (protect-advice.org.uk), the free whistleblowing advice line.

SUPPORT NCG

NCG does this work without vendor money, so our findings stay ours. If it has been useful to you, you can support NCG's independent work here. NCG is a for-profit company, not a charity. Contributions are not tax-deductible and buy no service or influence.

Come citare

Grizzle, H. M. (2026, August 17). L'IA può tradurre. Ma chi traduce il rischio?. Novara Consulting Group. https://doi.org/10.5281/zenodo.23003333