This document was translated from English by machine and may contain mistakes. The English original is at https://www.novaracg.com/2026/09/14/the-ai-ceos-found-the-brakes-sign-language-ai-should-pay-attention/.

I CEO dell'IA hanno trovato i freni... L'IA per la lingua dei segni dovrebbe farci caso

Heather Grizzle · September 14, 2026

Frontier AI CEOs are now asking for testing, audits and incident reporting. Sign language AI should hold itself to the same brakes.

Contenuti

1. La frontiera chiede freni
2. Il debito di governance è reale
3. Ciò che la frontiera sta imparando, l'IA per la lingua dei segni può impararlo ora
4. "We tested it" is not governance
5. Non aspettate il vostro momento di rallentamento
6. La garanzia indipendente non è un attacco all'innovazione
7. C'è ancora l'opportunità di anticipare questo cambiamento
8. L'avvertimento

La frontiera chiede freni

Per anni, la narrazione dominante nell'intelligenza artificiale è stata la velocità. Costruire il modello più capace. Lanciare l'agente. Battere il concorrente. Raccogliere capitali. Espandere il benchmark. Portare il prodotto sul mercato. La governance avrebbe potuto recuperare il terreno perduto in seguito.

Sembra che quel «in seguito» sia arrivato.

On September 9, OpenAI chief global affairs officer Chris Lehane published a call for mandatory, capability-based national AI safety regulation covering testing standards,

independent assessments, cybersecurity protections and incident reporting, alongside the company's endorsement of four California bills, including one establishing infrastructure for independent safety assessments and another setting standards for AI auditors.[^1] Three days later, Anthropic CEO Dario Amodei went considerably further. In an essay arguing that the industry must deliberately pace frontier development, he proposed embedding third-party evaluators inside frontier labs with near-employee access so they could independently examine safety measures, incidents and model behavior.

Poi è successo qualcosa di insolito.

Within hours, OpenAI CEO Sam Altman publicly agreed that the frontier needed pacing, and told Fortune days later that his company would not pursue a public offering this year because it had too much safety work left to do. Elon Musk, who had pointedly declined to join earlier industry calls for restraint, quoted the essay approvingly. Demis Hassabis, now chair of Google DeepMind and chief scientist of Alphabet, wrote that the essay pointed toward the right path forward. Companies locked in one of the most consequential technology races in history converged, inside a single day, on the proposition that capability could no longer be allowed to outrun control. Axios ran the moment under a headline that needed no elaboration: "AI's most powerful CEOs hit the brakes."

Today, September 14, Microsoft added another piece. Microsoft AI CEO Mustafa Suleyman released a draft code of conduct that he describes as a constitution of sorts for the company's future models. The draft would require Microsoft's systems to communicate clearly with humans, respect human boundaries, remain controllable, accept correction and shutdown, and treat behavior that violates the code as a model failure rather than an acceptable expression of autonomous judgment. Microsoft assembled it after expert consultations and has opened it to roughly six weeks of public feedback.[^3] Suleyman called it a warning shot.

Quella sequenza dovrebbe essere lettura obbligatoria per ogni azienda che sviluppa IA per la lingua dei segni. Non perché un avatar segnante stia per conquistare internet. Perché è così che si presenta il debito di governance.



Il debito di governance è reale

Le aziende tecnologiche comprendono il debito tecnico. Costruire abbastanza velocemente, rimandare abbastanza decisioni architetturali, e alla fine le scorciatoie diventano costose.

Il debito di governance funziona allo stesso modo, e si accumula attraverso decisioni che, singolarmente, sembrano ciascuna ragionevole. Un'azienda lancia un prodotto prima di definire chi è responsabile quando il sistema fallisce. Pubblica un'affermazione sull'accuratezza prima di stabilire cosa conta come errore. Costruisce un benchmark prima di chiedersi se quel benchmark rappresenti le persone che useranno effettivamente il prodotto. Convalida internamente perché la valutazione indipendente è scomoda. Aggiunge una supervisione umana senza specificare cosa la persona debba notare, quando debba intervenire, o se possieda affatto la competenza per riconoscere un fallimento. Nomina un gruppo consultivo dopo che le decisioni importanti sul prodotto sono già state prese. Definisce un sistema accessibile perché produce segni.

Ognuna di queste scelte crea debito di governance. E a differenza del debito tecnico, il costo non ricade solo sullo sviluppatore. Lo sostengono gli utenti. Lo sostengono gli acquirenti. Lo sostengono le agenzie pubbliche. Nella tecnologia della lingua dei segni, lo sostengono le persone Deaf.

Ciò che la frontiera sta imparando, l'IA per la lingua dei segni può impararlo ora

C'è un fatto straordinario al centro dell'attuale dibattito sull'IA. Le aziende con alcuni dei più grandi team di ingegneria, budget di ricerca e organizzazioni per la sicurezza al mondo stanno chiedendo maggiore controllo esterno. Anthropic propone una valutazione integrata da parte di terzi. OpenAI sostiene una regolamentazione federale obbligatoria e una valutazione indipendente. Microsoft sta cercando di formalizzare i confini del controllo umano. Dirigenti rivali discutono pubblicamente di contenimento coordinato, cosa che fino a poco tempo fa sarebbe stata impensabile.

La lezione non è che le aziende dovrebbero aspettare che un governo scriva uno standard perfetto. È quasi l'opposto. Le organizzazioni più vicine alla tecnologia stanno ammettendo che la fiducia interna non è prova sufficiente.

Questo conta per l'IA della lingua dei segni perché il settore è ancora abbastanza giovane da poter scegliere una sequenza diversa. La governance non deve essere il lavoro di riparazione. Può far parte dell'architettura.

"We tested it" is not governance

L'IA per la lingua dei segni ha una versione particolarmente pericolosa del problema della convalida. Un sistema può produrre qualcosa di riconoscibilmente simile a un segno e comunque fallire la persona che ne dipende. Una traduzione può essere fluente e sbagliata. Un avatar può essere visivamente curato pur introducendo distorsioni linguistiche. Un benchmark può riportare un numero impressionante pur rappresentando un dataset ristretto, un ambiente di segnazione limitato, o un compito che poco assomiglia alla realtà d'uso. Uno studio può includere una convalida umana pur basandosi su un campione troppo piccolo per sostenere l'affermazione commerciale poi ad esso associata. E una dimostrazione di prodotto può funzionare magnificamente senza stabilire che il sistema sia adatto alla sanità, all'istruzione, all'occupazione, alla comunicazione d'emergenza o ai servizi governativi.

Non si tratta semplicemente di questioni di apprendimento automatico. Sono questioni di garanzia. Chi l'ha testato, e chi ha selezionato il test? Quali segnanti erano rappresentati, e quali no? Cosa è stato misurato, e cosa è stato escluso? Cosa conta come errore rilevante? Quali contesti di implementazione sono stati valutati? Chi ha l'autorità di fermare l'implementazione? Cosa succede quando le prestazioni peggiorano, e chi esamina gli incidenti quando ciò accade? E, per una tecnologia costruita intorno alla comunicazione Deaf, dove entra effettivamente l'autorità Deaf nella decisione?

Queste domande diventano considerevolmente più difficili a cui rispondere dopo che i contratti sono stati firmati, le affermazioni di marketing pubblicate e ai clienti è stato detto che la tecnologia è pronta.

Non aspettate il vostro momento di rallentamento

There is an understandable temptation for smaller AI companies to look at the frontier governance debate and file it under somebody else's problem. Different technology, different scale, different risks. All correct. But risk magnitude is not the relevant comparison. Governance timing is.

L'industria di frontiera ha corso in avanti sotto una pressione competitiva straordinaria, e i suoi leader stanno ora affrontando domande sulla valutazione indipendente, i test obbligatori, la segnalazione degli incidenti, il controllo umano, la supervisione governativa e se il settore possa essere ritenuto affidabile nel giudicare quando i propri sistemi sono abbastanza sicuri. Gli sviluppatori di IA per la lingua dei segni possono osservare questo svolgersi in tempo reale, e poi decidere se ripeterlo.

Un'azienda che oggi sviluppa tecnologia per la lingua dei segni dovrebbe già essere in grado di rispondere a cinque domande.

1. Quale prova ci indurrebbe a non implementare il nostro stesso prodotto?
2. Chi, al di fuori dell'azienda, è autorizzato a contestare le nostre affermazioni?
3. Quale autorità detengono le persone Deaf al di là del fornire riscontri o dati di addestramento?
4. Quali usi della nostra tecnologia consideriamo inappropriati anche se un cliente fosse disposto a pagarli?
5. Cosa succede quando le prove post-implementazione contraddicono le prove che abbiamo usato per vendere il sistema?

Se queste risposte non esistono, l'azienda non ha semplicemente una documentazione di governance incompleta. Ha un debito di governance.

La garanzia indipendente non è un attacco all'innovazione

Qui è dove il mercato dell'IA per la lingua dei segni deve maturare rapidamente. Il controllo indipendente non è ostilità verso la tecnologia. È un segnale che la tecnologia conta abbastanza da essere esaminata con attenzione.

I produttori di aeromobili non stabiliscono la fiducia invitando gli acquirenti ad ammirare l'aereo. I produttori di dispositivi medici non stabiliscono la sicurezza pubblicando testimonianze. Le revisioni contabili non permettono che l'entusiasmo del management sostituisca le prove. Man mano che l'IA entra nell'accessibilità, si applica la stessa distinzione. Uno sviluppatore crea il sistema. La governance stabilisce le regole che lo circondano. La valutazione mette alla prova le prove. La garanzia si chiede se quelle prove siano sufficienti a sostenere l'affermazione fatta. Sono quattro funzioni

diverse, e quando un'unica azienda le controlla tutte e quattro, gli acquirenti dovrebbero porsi delle domande.

The frontier industry is beginning to acknowledge exactly that problem. Anthropic's proposal for embedded outside evaluators is significant precisely because it moves independent scrutiny closer to the development process rather than treating evaluation as a ceremonial checkpoint immediately before release.

L'IA per la lingua dei segni dovrebbe andare ancora oltre. La valutazione tecnica indipendente è necessaria, ma la legittimità linguistica, il contesto d'implementazione, l'impatto sull'accessibilità e la governance Deaf non possono essere ridotti a test ingegneristici. Un modello può funzionare esattamente come progettato e comunque essere inadatto al compito per cui è stato venduto.

C'è ancora l'opportunità di anticipare questo cambiamento

Questo è ciò che le aziende di IA per la lingua dei segni non dovrebbero lasciarsi sfuggire. Stanno ricevendo un avvertimento insolitamente prezioso, e non devono necessariamente vivere prima la crisi di governance.

Il lavoro è specifico. Documentare i limiti prima che il marketing li aggiri con la scrittura. Separare le prestazioni dimostrative dalle prove d'implementazione. Usare valutatori indipendenti prima che un cliente li richieda. Creare procedure per gli incidenti prima del primo incidente grave. Stabilire i contesti in cui il prodotto non verrà venduto. Dare agli esperti Deaf una reale autorità decisionale anziché una presenza puramente consultiva. Rendere le affermazioni di convalida abbastanza specifiche da poter essere esaminate e contestate da un'altra parte. Conservare le prove in modo che un acquirente possa in seguito determinare perché una decisione d'implementazione fosse giustificata al momento in cui è stata presa.

Ecco cosa significa diventare garantibili. Ed è anche considerevolmente più economico che ricostruire la fiducia in seguito.

L'avvertimento

Le aziende di IA più potenti al mondo stanno scoprendo che la governance non può restare indietro rispetto alla capacità per sempre. Alcuni dei loro leader stanno ora chiedendo freni. Altri chiedono leggi. Altri ancora stanno redigendo costituzioni, strutture di valutazione esterna e nuovi meccanismi di controllo. Dispongono di risor-

se che la maggior parte delle aziende di tecnologia per l'accessibilità non possederà mai, eppure anche loro stanno scoprendo quanto diventi difficile la governance una volta che la tecnologia sta già viaggiando a piena velocità.

Sign language AI companies should study that carefully. Do not wait until a failed deployment, a procurement challenge, a regulator, a lawsuit or the Deaf community forces the governance conversation. Do not build the evidence after making the claim. Do not invite independent scrutiny only after trust has been lost. And do not confuse being first to market with being ready for the responsibility that comes with entering someone else's language, communication and access.

L'IA di frontiera sta ora cercando di trovare i propri freni. L'IA per la lingua dei segni ha ancora l'opportunità di installarli prima di accelerare.

NOVARA CONSULTING GROUP

Novara Consulting Group fornisce assicurazione sull'IA, governance dell'accessibilità e valutazione indipendente.

FOOTNOTES:
Chris Lehane, "The AI policy window is open. We need to act.," OpenAI, September 9, 2026; Reuters, "OpenAI pushes for mandatory national AI safety requirements," September 9, 2026.
Axios, "AI's most powerful CEOs hit the brakes," September 13, 2026.
Jeffrey Dastin, "Microsoft drafts code of conduct to keep its AI under human control," Reuters, September 14, 2026.

> **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, September 14). I CEO dell'IA hanno trovato i freni… L'IA per la lingua dei segni dovrebbe farci caso. Novara Consulting Group. https://doi.org/10.5281/zenodo.23003722