NOVARA
CONSULTING GROUP

The AI CEOs Found the Brakes... Sign Language AI Should Pay Attention

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.

Contents

1. The frontier is asking for brakes
2. Governance debt is real
3. What the frontier is learning, sign language AI can learn now
4. "We tested it" is not governance
5. Do not wait for your own slow-down moment
6. Independent assurance is not an attack on innovation
7. There is still an opportunity to get ahead of this
8. The warning

The frontier is asking for brakes

For years, the dominant story in artificial intelligence was speed. Build the more capable model. Ship the agent. Beat the competitor. Raise the capital. Expand the benchmark. Get the product into the market. Governance could catch up later.

Later appears to have arrived.

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.

Then something unusual happened.

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.

That sequence should be required reading for every company developing sign language AI. Not because a signing avatar is about to take over the internet. Because this is what governance debt looks like.

Governance debt is real

Technology companies understand technical debt. Build quickly enough, defer enough architecture decisions, and eventually the shortcuts become expensive.

Governance debt works the same way, and it accumulates through decisions that each look reasonable in isolation. A company launches before defining who is accountable when the system fails. It publishes an accuracy claim before establishing what counts as an error. It builds a benchmark before asking whether that benchmark represents the people who will actually use the product. It validates internally because indepen-

dent evaluation is inconvenient. It adds human oversight without specifying what the human is supposed to notice, when that person must intervene, or whether they hold the expertise to recognize failure at all. It appoints an advisory group after the important product decisions have already been made. It calls a system accessible because the system produces signs.

Every one of those choices creates governance debt. And unlike technical debt, the cost is not carried only by the developer. Users carry it. Purchasers carry it. Public agencies carry it. In sign language technology, Deaf people carry it.

What the frontier is learning, sign language AI can learn now

There is an extraordinary fact sitting inside the current AI debate. The companies with some of the largest engineering teams, research budgets and safety organizations in the world are asking for more outside scrutiny. Anthropic is proposing embedded third-party evaluation. OpenAI is backing mandatory federal regulation and independent assessment. Microsoft is attempting to formalize the boundaries of human control. Rival executives are discussing coordinated restraint in public, which until recently would have been unthinkable.

The lesson is not that companies should wait until a government writes a perfect standard. It is closer to the opposite. The organizations nearest to the technology are conceding that internal confidence is not sufficient evidence.

That matters for sign language AI because the field is still young enough to choose a different sequence. Governance does not have to be the repair job. It can be part of the architecture.

"We tested it" is not governance

Sign language AI has a particularly dangerous version of the validation problem. A system can produce something recognizably sign-like and still fail the person depending on it. A translation can be fluent and wrong. An avatar can be visually polished while introducing linguistic distortion. A benchmark can report an impressive number while representing a narrow dataset, a constrained signing environment, or a task bearing little resemblance to deployment. A study can include human validation while relying on a sample far too small to support the commercial claim eventually attached

to it. And a product demonstration can work beautifully without establishing that the system is appropriate for healthcare, education, employment, emergency communication or government services.

These are not simply machine learning questions. They are assurance questions. Who tested it, and who selected the test? Which signers were represented, and which were not? What was measured, and what was excluded? What counts as a material error? Which deployment contexts were evaluated? Who holds the authority to stop deployment? What happens when performance degrades, and who reviews incidents when it does? And, for technology built around Deaf communication, where does Deaf authority actually enter the decision?

Those questions become considerably harder to answer after contracts are signed, marketing claims are published and customers have been told the technology is ready.

Do not wait for your own slow-down moment

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.

The frontier industry raced forward under extraordinary competitive pressure, and its leaders are now confronting questions about independent evaluation, mandatory testing, incident reporting, human control, government oversight and whether the industry can be trusted to judge when its own systems are safe enough. Sign language AI developers can watch that unfold in real time, and then decide whether to repeat it.

A company developing sign language technology today should already be able to answer five questions.

1. What evidence would cause us not to deploy our own product?
2. Who outside the company is permitted to challenge our claims?
3. What authority do Deaf people hold beyond providing feedback or training data?
4. Which uses of our technology do we consider inappropriate even if a customer is willing to pay for them?

5. What happens when post-deployment evidence contradicts the evidence we used to sell the system?

If those answers do not exist, the company does not merely have unfinished governance documentation. It has governance debt.

Independent assurance is not an attack on innovation

This is where the sign language AI market needs to mature quickly. Independent scrutiny is not hostility toward the technology. It is a signal that the technology matters enough to scrutinize.

Aircraft manufacturers do not establish confidence by inviting purchasers to admire the aircraft. Medical device manufacturers do not establish safety by publishing testimonials. Financial audits do not permit management enthusiasm to substitute for evidence. As AI moves into accessibility, the same distinction applies. A developer creates the system. Governance establishes the rules surrounding it. Evaluation tests the evidence. Assurance asks whether that evidence is sufficient to support the claim being made. Those are four different functions, and when a single company controls all four, buyers should ask questions.

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.

Sign language AI should go further still. Independent technical evaluation is necessary, but linguistic legitimacy, deployment context, accessibility impact and Deaf governance cannot be reduced to engineering tests. A model can function exactly as designed and still be unsuitable for the job it has been sold to perform.

There is still an opportunity to get ahead of this

This is the part sign language AI companies should not miss. They are being handed an unusually valuable warning, and they do not have to experience the governance crisis first.

The work is specific. Document limitations before marketing writes around them. Separate demonstration performance from deployment evidence. Use independent evalu-

ators before a customer demands them. Create incident procedures before the first serious incident. Establish the contexts in which the product will not be sold. Give Deaf experts actual decision authority rather than consultative presence. Make validation claims specific enough that another party can examine and challenge them. Preserve evidence so a purchaser can later determine why a deployment decision was justified at the time it was made.

That is what becoming assurable looks like. It is also considerably cheaper than rebuilding trust later.

The warning

The most powerful AI companies in the world are discovering that governance cannot permanently remain behind capability. Some of their leaders are now asking for brakes. Others are asking for laws. Others are drafting constitutions, external evaluation structures and new control mechanisms. They command resources that most accessibility technology companies will never possess, and even they are finding out how difficult governance becomes once the technology is already moving at full speed.

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.

Frontier AI is now trying to find its brakes. Sign language AI still has the opportunity to install them before accelerating.

NOVARA CONSULTING GROUP

Novara Consulting Group provides AI assurance, accessibility governance and independent evaluation.

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.

How to cite

Grizzle, H. M. (2026, September 14). The AI CEOs Found the Brakes… Sign Language AI Should Pay Attention. Novara Consulting Group. https://doi.org/10.5281/zenodo.23003722