NOVARA
CONSULTING GROUP

The Policy Problem With Sign Language AI When Public Procurement Outruns Scientific Evidence

Heather Grizzle · July 7, 2026

Governments are buying sign language AI before independent evidence exists. Why procurement is the checkpoint and what responsible buying looks like.

Contents

1. The Procurement Moment
2. The Appeal of Technological Promise
3. The Evidentiary Scarcity Problem
4. Why Sign Language AI Is Not Just Another Accessibility Tool
5. Scientific Inquiry vs. Persuasive Writing
6. Procurement as a Governance Checkpoint
7. What Responsible Evaluation Should Ask
8. The Risk of Repeating the Hype Cycle
9. The Role of Independent Evaluation and Governance
10. Evidence Before Adoption

Governments are purchasing language access systems before the evidence base has matured. This is a public accountability crisis in slow motion.

The Procurement Moment

Government agencies are being asked to make significant purchasing decisions about sign language AI systems before the scientific evidence base has matured. This is not a question of whether the technology will eventually mature. It is a question of whether current vendor claims are being treated as validated fact when they remain unverified promise. The distinction matters, especially when public money and Deaf community access are at stake.

Over the past eighteen months, accessibility officers, procurement teams, and IT departments across federal, state, and municipal governments have begun receiving vendor pitches for sign language AI systems. These pitches emphasize innovation, efficiency, and inclusion. They promise to expand access at scale, reduce costs, and modernize language services. The messaging is compelling, and it arrives at an institutional moment when public sector budgets are constrained and accessibility mandates are expanding. The question being asked by procurement officers is simple: "Why wouldn't we adopt this?"

The answer is equally simple, but less frequently voiced: "Because we do not yet have independent evidence that these systems work as claimed, and we have no agreed standards for what 'working' means in the context of language access."

This is not a technical objection. It is a governance objection. It is the question of whether public institutions should normalize purchasing decisions that lack the accountability structures necessary to protect the people most affected by those decisions.

The Appeal of Technological Promise

Sign language AI is attractive to public institutions for reasons that are entirely rational within the logic of institutional procurement. The technology promises innovation in a space where innovation has been episodically limited. It suggests scalability in contexts where language services have historically required one-to-one human interpreters. It offers cost efficiency in systems under budget pressure. It frames itself as inclusion and empowerment, aligned with disability rights and accessibility mandates.

These are powerful procurement incentives. When an organization faces simultaneous pressure to expand access, reduce costs, and adopt emerging technologies, a solution that promises all three becomes difficult to refuse at the moment of the pitch.

However, procurement convenience is not the same as civil rights compliance. Institutional efficiency is not the same as linguistic validity. Cost savings are not the same as community safety. The appeal of technological promise has repeatedly driven public sector adoption of systems before their real-world effects have

Procurement convenience is not the same as civil rights compliance.

been fully understood. In accessibility, this pattern is especially dangerous because the people most likely to experience harm are rarely the people deciding whether to adopt the technology.

The Evidentiary Scarcity Problem

Many sign language AI systems are promoted with claims about accessibility, linguistic competence, inclusion, language access, and Deaf empowerment. These are claims that matter. They address real needs. But these claims are often not accompanied by the evidence structures that would normally be required before a public institution normalizes a system affecting fundamental aspects of people's lives.

Specifically, these systems are frequently promoted without:

- independent evaluation by researchers outside the vendor organization;
- transparent and publicly available datasets;
- published error analysis that documents what the system fails to do and under what conditions;
- real-world deployment evidence from sustained use in operational contexts; or
- community-informed risk assessment conducted with Deaf native signers and the populations most dependent on language access.

This is not to say that all sign language AI vendors are making unfounded claims. It is to say that even when claims are made in good faith, the absence of independent verification means that institutional decisions are being made on the basis of persuasive writing rather than scientific evidence. Those are not the same thing. Vendor demonstrations, testimonials from trial users, case studies, and marketing materials can all be compelling and can all fall short of the evidence standard necessary for responsible public procurement.

The gap is not trivial. It is the space between "this technology shows promise" and "this technology is ready for public deployment." That space is where public policy should insist on accountability.

Why Sign Language AI Is Not Just Another Accessibility Tool

Sign language is a full natural language, not a feature, caption layer, or customer service enhancement. It has complex phonological structure, syntactic rules, morphological processes, and pragmatic norms. It is the native language of Deaf communities and the foundation of educational, professional, and civic participation for millions of people globally.

This matters because errors in sign language AI can affect more than convenience. They can affect comprehension of information that affects health, legal status, educational access, employment eligibility, emergency response, and civic rights. A captioning system that occasionally misses dialogue is an accessibility feature with a known limitation. A sign language system that occasionally misrepresents meaning is a potential vector for harm.

The error tolerance for sign language AI is consequently different from the error tolerance for less linguistically fundamental technologies. A system that works correctly 95 percent of the time might be acceptable for a supplementary feature. It may not be acceptable as the primary language access mechanism in a legal proceeding, medical consultation, or educational setting. Public agencies need to understand not only what a system does well, but what it does poorly and whether those failures can be tolerated in the specific context where the system will be deployed.

Scientific Inquiry vs. Persuasive Writing

The distinction between scientific inquiry and persuasive writing is not academic. It is a matter of public accountability.

Vendor decks, product demonstrations, testimonials, case studies, and marketing language serve important communicative functions. They tell coherent stories. They emphasize benefits. They create momentum. But they are persuasive tools, not scientific evidence. Scientific evidence is produced through processes designed to be inspected, questioned, replicated, and challenged. It involves independent evaluation. It documents uncertainty and failure modes, not just successes. It is subject to peer review and public scrutiny.

When public agencies make procurement decisions, they are implicitly deciding which form of knowledge they trust. If they trust vendor persuasion more than independent evidence, they are making a governance decision with consequences. They are saying that institutional efficiency and innovation momentum are more important than independent verification of claims that affect public accountability.

This is not an argument against innovation. It is an argument for the discipline that should accompany innovation in public institutions. Government procurement exists to serve the public. It requires evidence, not just promise.

Procurement as a Governance Checkpoint

Procurement is one of the most important intervention points in technology governance, and it is frequently treated as a transaction rather than a decision.

Once a system is purchased, integrated into institutional workflows, trained into staff processes, and normalized as standard operating procedure, changing course becomes difficult. The installed base creates inertia. The sunk costs create resistance to alternatives. The staff who have learned to work around the system′s limitations develop workarounds rather than escalate problems. The system becomes infrastructure, and infrastructure is hard to change.

This is why procurement must be treated as a governance checkpoint. It is the moment where a public agency has leverage to demand accountability. It is the moment when vendors are most responsive to requirements and most willing to demonstrate evidence. It is the moment before the system becomes normalized and difficult to evaluate critically.

Responsible procurement should therefore require independent evidence before adoption, not require apologies after deployment. The question should not be ″How do we make this system work?″ but rather ″Do we have sufficient evidence that this system is appropriate for this use case?″ The burden of proof should rest with the vendor and the evidence, not with the institution and hope.



What Responsible Evaluation Should Ask

Institutions approaching sign language AI procurement should ask clear, specific questions before making purchasing decisions. These questions should focus on evidence, transparency, community engagement, risk assessment, and governance.

- Who conducted the evaluation of this system, and were they independent of the vendor?
- What sign languages, dialects, registers, and user populations have been tested, and what populations were excluded from testing?
- Were Deaf native signers involved in evaluation, and in what capacity?
- Have DeafBlind users, disabled users, and multiply marginalized users been specifically tested?
- What are the documented failure modes, and under what conditions does the system perform poorly?
- What happens in high-stakes contexts such as medical communication, legal proceedings, or education?
- Can the vendor articulate clearly what the system should not be used for, based on evidence?
- Does the system disclose its own uncertainty, or does it present outputs with false confidence?
- Is there a human override or escalation process built into deployment?
- Can the public agency audit system performance independently after purchase, or is performance evaluation limited to vendor-provided metrics?

These are not hostile questions. They are the questions that responsible institutions should ask before making decisions that affect fundamental language access for Deaf communities.

The Risk of Repeating the Hype Cycle

Public institutions have repeatedly adopted technologies before fully understanding downstream effects. Voice recognition systems were normalized before error rates for accented speech were fully documented. Predictive algorithms were deployed in criminal justice before bias in training data was systematically studied. Automated cap-



tioning systems were adopted before their limitations for complex dialogue and specialized domains were mapped.

In accessibility specifically, this pattern is especially consequential. The people most affected by accessibility technology decisions are rarely the people making those decisions. Procurement officers, IT directors, budget officers, and executives make the purchasing call. Users from the affected community often discover the technology's limitations only after deployment, when the purchasing decision has already been locked in and alternatives are no longer being considered.

The risk of repeating this cycle with sign language AI is real. The momentum is building. The vendor pitches are becoming more polished. The pressure to adopt is increasing. And the evidence base, while growing, is still not mature enough to support the scale of deployment currently being proposed.

The Role of Independent Evaluation and Governance

Novara Consulting Group was founded to address a visible gap in procurement practice. Public institutions lack a structured approach to evaluating sign language AI systems before adoption, and that gap creates risk for both the institution and the communities most dependent on language access.

The Sign Language Access Trust (SLAT) Index is a governance-focused framework designed to help institutions ask better questions before purchasing sign language AI systems. It is not a technical standard, and it is not a certification system. It is a structured evaluation approach that focuses on evidence transparency, governance mechanisms, community engagement, linguistic risk assessment, deployment context, and accountability structures.

SLAT assesses systems across ten governance domains:

> **THE TEN SLAT GOVERNANCE DOMAINS**
>
> product transparency, evaluation independence, dataset documentation, error analysis, real-world validation, user engagement, deployment safeguards, escalation and override, performance monitoring, and community accountability

For each domain, SLAT asks whether vendors can provide evidence that meets standard criteria, or whether gaps remain unverified.

The goal is not to block innovation. It is to make sure that the questions being asked before purchase are the right questions, and that the answers are being subjected to independent scrutiny.

Evidence Before Adoption

The future of sign language AI should not be decided by marketing momentum alone. If public agencies are going to purchase systems that mediate language access for Deaf communities, they must demand independent evidence before adoption. That evidence should be inspected before commitment, not debated after deployment. The question is not whether sign language AI is promising. The question is whether public institutions are prepared to govern that promise responsibly.

The distinction between innovation and accountability is not a constraint on progress. It is a prerequisite for progress that is safe. Public agencies that treat procurement as a governance checkpoint, not a transaction, will find vendors and systems that are willing to meet that standard. The ones that do will set a precedent for the field. The ones that don't will bear the consequences.

ABOUT THE AUTHOR

Heather M. Grizzle-Odland

Author Note: Heather M. Grizzle-Odland is Founder and CEO of Novara Consulting Group LLC, an AI governance and accessibility consulting firm. She is the developer of the Sign Language Access Trust (SLAT) Index and a doctoral student in Public Policy at Liberty University's Helms School of Government. She writes from inside the Deaf community and is a non-speaking ASL user based in Minnesota.

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, July 7). The Policy Problem With Sign Language AI When Public Procurement Outruns Scientific Evidence. Novara Consulting Group. https://www.novaracg.com/2026/07/07/the-policy-problem-with-sign-language-ai-when-public-procurement-outruns-scientific-evidence/