NOVARA
CONSULTING GROUP

This document was translated from English by machine and may contain mistakes. The English original is at https://www.novaracg.com/2026/07/07/the-policy-problem-with-sign-language-ai-when-public-procurement-outruns-scientific-evidence/.

公共采购超越科学证据时,手语人工智能面临的政策问题

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.

目录

1. 采购时刻
2. 技术承诺的吸引力
3. 证据稀缺问题
4. 为何手语人工智能不只是又一种无障碍工具
5. 科学探究与说服性文字的区别
6. 采购作为治理关卡
7. 负责任的评估应当追问什么
8. 重蹈炒作周期覆辙的风险
9. 独立评估与治理的作用
10. 证据先于采纳

政府部门正在证据基础尚未成熟之前购置语言无障碍系统。这是一场正在慢镜头中上演的公共问责危机。

采购时刻

各级政府机构正被要求在科学证据基础尚未成熟之前,就手语人工智能系统做出重大采购决策。这并非该技术最终是否会成熟的问题,而是当前供应商的宣称是否被当作已验

证的事实来对待,尽管它们实际上仍只是未经证实的承诺。这一区别至关重要,尤其是当公共资金与聋人社群的无障碍权益处于风险之中时。

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."

这不是一个技术层面的异议,而是一个治理层面的异议——即公共机构是否应当将缺乏必要问责结构的采购决策常态化,而这些问责结构本应用来保护受这些决策影响最深的人群。

技术承诺的吸引力

从公共机构采购逻辑的角度来看,手语人工智能之所以具有吸引力,理由完全合乎理性。该技术承诺在一个创新历来有限的领域带来突破;它暗示在历来依赖一对一人工传译的语境中实现可扩展性;它在预算压力下的系统中提供成本效益;它以包容与赋权自我定位,并与残障权利及无障碍要求相契合。

这些都是强有力的采购诱因。当一个组织同时面临扩大服务覆盖、降低成本、采用新兴技术的压力时,一个承诺三者兼得的解决方案,在推介的那一刻很难被拒绝。

然而,采购上的便利并不等同于民权合规。机构效率并不等同于语言学有效性。成本节约并不等同于社群安全。技术承诺的吸引力一再驱使公共部门在充分了解其现实影响之前就采纳相关系统。在无障碍领域,这种模式尤为危险,因为最有可能受到伤害的人群,往往不是那些决定是否采用该技术的人。

> 采购上的便利,不等于民权合规。

证据稀缺问题

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.

具体而言,这些系统的推广往往缺少以下要素:

- 由供应商组织之外的研究人员进行的独立评估;
- 透明且可公开获取的数据集;
- 记录系统在何种条件下无法完成何种任务的已发表错误分析;
- 来自实际运营环境中持续使用的现实部署证据;或
- 与聋人母语手语使用者及最依赖语言无障碍服务的群体共同开展的、以社群为本的风险评估。

这并不是说所有手语人工智能供应商的宣称都毫无根据。而是说,即便这些宣称是出于善意提出的,缺乏独立验证也意味着机构决策是建立在有说服力的文字表述之上,而非科学证据之上。这两者并不相同。供应商演示、试用用户的证言、案例研究和营销材料都可能极具说服力,但也都可能达不到负责任的公共采购所需的证据标准。

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.

为何手语人工智能不只是又一种无障碍工具

手语是一门完整的自然语言,而不是一项功能、一层字幕或客服增强手段。它拥有复杂的音系结构、句法规则、构词过程和语用规范。它是聋人社群的母语,也是全球数以百万计人口教育、职业和公民参与的基础。

这一点之所以重要,是因为手语人工智能中的错误所影响的可能远不止便利性。它们可能影响到与健康、法律身份、教育机会、就业资格、紧急响应和公民权利相关的信息理解。一个偶尔漏译对话的字幕系统,是一项存在已知局限的无障碍功能;而一个偶尔曲解含义的手语系统,则可能是造成伤害的潜在途径。

因此,手语人工智能的容错标准与那些语言学重要性较低的技术不同。一个准确率达到95%的系统,或许可以作为辅助功能被接受;但作为法律程序、医疗咨询或教育场景中主要的语言无障碍机制,这一准确率可能是不可接受的。公共机构不仅需要了解某系统擅长什么,还需要了解它在哪些方面表现不佳,以及在该系统将被部署的具体场景中,这些失误是否可以被容忍。

科学探究与说服性文字的区别

科学探究与说服性文字之间的区别,不是学术层面的问题,而是公共问责层面的问题。

供应商的演示文稿、产品演示、证言、案例研究和营销语言都发挥着重要的传播功能。它们讲述连贯的故事,强调好处,营造势头。但它们是说服工具,而不是科学证据。科学证据是通过旨在被审查、质疑、复现和挑战的过程产生的。它涉及独立评估,记录不确定性和失败模式,而不仅仅是成功案例。它需要接受同行评审和公众监督。

当公共机构做出采购决策时,实际上是在隐含地决定他们信任哪种知识形式。如果他们更信任供应商的说服而非独立证据,那么他们做出的就是一项带有后果的治理决策。这意味着他们在宣告:机构效率和创新势头,比对影响公共问责的宣称进行独立验证更为重要。

这不是反对创新的论点,而是主张公共机构在推进创新时应当保持的纪律性。政府采购的存在是为了服务公众,它需要的是证据,而不仅仅是承诺。

采购作为治理关卡

采购是技术治理中最重要的干预节点之一,但它常常被当作一笔交易,而非一项决策来对待。

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.

这正是为什么必须把采购当作一个治理关卡来对待。这是公共机构有筹码要求问责的时刻;是供应商对各项要求最为积极响应、也最愿意展示证据的时刻;也是系统尚未被常态化、尚可被批判性评估的最后时机。

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.

负责任的评估应当追问什么

在着手手语人工智能采购的机构,应当在做出采购决策之前提出清晰、具体的问题。这些问题应聚焦于证据、透明度、社群参与、风险评估和治理。

- 是谁对该系统进行了评估,他们是否独立于供应商?
- 该系统测试过哪些手语、方言、语域和用户群体,又有哪些群体被排除在测试之外?
- 聋人母语手语使用者是否参与了评估,他们以何种身份参与?
- 聋盲用户、残障用户以及多重边缘化用户是否已被专门测试过?
- 已记录在案的失效模式有哪些?系统在何种条件下表现不佳?
- 在医疗沟通、法律诉讼或教育等高风险场景中会发生什么?
- 供应商能否基于证据清楚说明该系统不应被用于哪些场景?
- 该系统是否会披露自身的不确定性,还是以虚假的自信呈现输出结果?
- 部署过程中是否内置了人工干预或升级处理机制?
- 公共机构在采购后能否独立审计系统表现,还是只能依赖供应商提供的指标进行评估?

这些并非带有敌意的提问。它们是负责任的机构在做出影响聋人群体基本语言获取权的决策之前,理应提出的问题。

重蹈炒作周期覆辙的风险

公共机构一再在尚未充分理解下游影响之前就采用新技术。语音识别系统在带口音语音的错误率被充分记录之前就已被常态化使用。预测算法在训练数据中的偏见被系统研究

之前就已被部署于刑事司法领域。自动字幕系统在其对复杂对话和专业领域的局限性被摸清之前就已被采用。

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.

手语人工智能重蹈这一循环的风险是真实存在的。势头正在积聚。供应商的推介说辞正变得愈发精致。采用的压力正在增加。而证据基础虽在不断积累,却仍不足以支撑当前所提议的部署规模。

独立评估与治理的作用

Novara Consulting Group 的成立,正是为了填补采购实践中一个显而易见的空白。公共机构缺乏一套结构化的方法,来在采用手语人工智能系统之前对其进行评估,而这一空白给机构本身以及最依赖语言获取权的群体都带来了风险。

Sign Language Access Trust (SLAT) Index 是一个以治理为核心的框架,旨在帮助机构在购买手语人工智能系统之前提出更好的问题。它不是一项技术标准,也不是一套认证体系。它是一种结构化的评估方法,聚焦于证据透明度、治理机制、社区参与、语言风险评估、部署情境以及问责结构。

SLAT 从十个治理维度对系统进行评估:

> **SLAT 的十个治理维度**
>
> 产品透明度、评估独立性、数据集文档、错误分析、真实世界验证、用户参与、部署保障措施、升级与干预机制、性能监测,以及社区问责

针对每一个维度,SLAT 都会询问供应商能否提供符合标准准则的证据,或者是否仍存在未经核实的空白。

其目的并非阻碍创新,而是确保在购买之前所提出的问题是正确的问题,并且这些答案要经受独立审查。

证据先于采纳

手语人工智能的未来不应仅由营销势头来决定。如果公共机构要购买那些为聋人群体语言获取权提供中介的系统,它们就必须在采用之前要求提供独立证据。这些证据应在承诺之前接受审查,而不是在部署之后才被讨论。问题不在于手语人工智能是否具有前景,而在于公共机构是否准备好负责任地治理这一前景。

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.

关于作者

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.

引用方式

Grizzle, H. M. (2026, July 7). 公共采购超越科学证据时,手语人工智能面临的政策问题. Novara Consulting Group. https://www.novaracg.com/zh/2026/07/07/the-policy-problem-with-sign-language-ai-when-public-procurement-outruns-scientific-evidence/