NOVARA
CONSULTING GROUP

This document was translated from English by machine and may contain mistakes. The English original is at https://www.novaracg.com/2026/07/22/the-procurement-theater-problem-procedural-assurance-evidentiary-assurance-and-the-governance-of-accessibility-ai/.

采购剧场问题:程序性保障、证据性保障与无障碍AI的治理

Heather Grizzle · July 22, 2026

Checklists, demos and VPATs create confidence without evidence. Procurement theater, and the evidentiary assurance accessibility AI buyers need.

目录

1. 执行摘要
2. 政策问题
3. 机构分析
4. 治理影响
5. NCG政策立场
6. 实施考量
7. 结语立场
8. 参考文献
9. 作者说明

执行摘要

为通信无障碍而采购人工智能系统的机构越来越自信地认为自己已经妥善管理了风险,而这种自信也越来越站不住脚。这种自信建立在各种文件之上:填写完成的核对清单、供应商文档、脚本化的演示、经过谈判的保修条款、备案的无障碍声明。每一份文件都是真实的。但无论单独看还是合起来看,它们都不是为了回答机构真正需要回答的问题而设计的,那个问题是:该系统在其被采购用于的部署环境中,是否能够安全、可靠、恰当地运行。

本期将这种状况称为“采购剧场”,并论证这是结构性问题,而非出于恶意。采购专业人员因程序上的完整性而受到嘉奖。供应商则因具有说服力的演示而获得回报。双方都可以完全履行各自的义务,而核心的性能问题却始终无人审视。补救之道不在于更多的文件材料,而在于将治理标准从“程序性保证”——询问是否采取了所要求的步骤,转向“证据性保证”——询问所收集的证据是否有能力发现真相。

Novara Consulting Group′s position is that this distinction will become the defining competency of public sector AI governance, and that institutions should begin building the evidentiary capacity to make it now, before deployment rather than after failure.

政策问题

治理的衡量标准不是活动本身,而是所进行的活动是否有能力发现真相。当前大多数无障碍人工智能采购未能通过这一检验,原因不在于流程草率,而在于这一流程本是为另一类产品设计的。

Traditional procurement instruments assume that a product′s capabilities are stable, observable, and largely disclosed by the manufacturer. Under that assumption, collecting artifacts is a reasonable proxy for verifying performance. Artificial intelligence systems mediating linguistic access violate every part of the assumption. Performance varies by signer, register, regional variation, lighting, camera geometry, latency, subject matter, and the presence or absence of the non-manual grammatical features from which this publication takes its name. A system that performs adequately in a controlled demonstration may degrade substantially in a clinical intake, a disciplinary hearing, or an emergency communication. Nothing in a standard artifact package is constructed to reveal that degradation.

三种混淆维持着这一问题。文档被当作验证,尽管文档记录的是供应商的主张,而非系统的实际表现。认证被当作性能,尽管认证通常证明的是流程的合规性,而非产出的质量。合规被当作有效性,尽管合规衡量的是与技术标准的一致性,而非聋人或听障用户是否实现了功能性的沟通无障碍。一个无法区分这三个概念的采购流程,所产生的机构风险与其填写的表格数量成正比,因为每份填写完成的表格都在增加信心,却并未增加知识。

机构分析

解释这种模式持续存在的原因,激励结构比任何“恶意论”都更有说服力。采购部门在审计制度下运作,该制度审查的是所要求的步骤是否已经执行,而非这些步骤在认识论上是否充分。一份包含供应商无障碍合规报告、演示记录以及经协商合同条款的档案能够通过审计。而一份记录了机构独立针对具有代表性的聋人用户、在具有代表性的条件下测试了该系统、并发现其存在不足的档案,则会带来进度风险和政治风险。这一体系奖励的是前者。

供应商面临的是对称的激励结构。无障碍人工智能市场奖励的是演示质量、叙事的连贯性以及签约速度。供应商很少被要求披露失败情形,而单方面披露这些情形会带来竞争上的代价。目前尚无公开证据支持“现有商用手语人工智能系统已在独立构建的、具有部署代表性的测试条件下接受评估并公布结果”这一说法。在缺乏此类证据的情况下,营销宣传和自行生成的文档便默认成为机构记录。

受影响的群体承担了剩余风险。聋人和听障用户通常不是采购的当事方,极少参与验收测试,只有在机构信心已经建立、合同已经签署之后,他们才会接触到该系统。到那时,举证责任已经发生倒置:用户被要求证明系统存在故障,而与之对立的是机构记录所声称的“已尽职调查、系统被认定合规”。这种倒置正是采购剧场造成的实质危害,而它完全是由遵循合法程序的合法行为者所造成的。

> 不去核查这一决定本身就是一项治理决定。

以政府工具传统的视角来理解,采购并非行政形式,而是一种公共行动工具,而工具的选择决定了谁有权发言、什么可以作为证明、以及问责最终落在何处。结合因果叙事文献来看,当前关于无障碍人工智能失败的主流叙述是一种“意外”的故事,即不可预测的技术偶尔表现不佳,而没有任何行为者需要承担责任。本文所倡导的证据性框架支持另一种解读——至少是“疏忽”的叙述:这些失败是可预见的,导致其发生的条件是可以事先发现的,而选择不去查看,本身就是一种治理决定。

治理影响

从程序性保证转向证据性保证,要求机构提出一套不同的问题,并将这些问题的答案视为验收标准,而非背景信息。所提出主张的支撑证据是什么,这些证据是性能证据还是流程证据?证据是由谁生成的,该方与结果之间存在何种商业关系,其方法透明度如何?该结果

能否由与采购结果无利害关系的一方独立复现?系统在何种具体条件下会失败,失败率是多少,在不同用户群体中严重程度的分布又如何?当这些失败在部署中发生时,由哪一方承担责任,通过何种机制、在何种时限内承担,受影响个人又能获得何种补救?

这些问题所做的远不止提高证据门槛,它们还重新分配了问责归属。一份能够回答这些问题的采购档案,确立了机构知道了什么、何时知道、以及机构接受了什么。而一份无法回答这些问题的采购档案,只能证明机构完成了一个流程。在新兴的治理预期之下——包括要求组织明确使用情境并记录剩余风险的基于风险的框架——后一种档案将越来越难以自圆其说。

此处所涉及的能力,是机构区分“信心”与“证据”的能力。这种能力并不存在于核对清单之中,而是存在于评估工具之中,存在于明确具有代表性的测试条件的能力之中,也存在于愿意将供应商的主张视为一种假设、而非一项结论的态度之中。

NCG政策立场

Novara Consulting Group主张,无障碍人工智能的采购应当由证据性保证来治理,以下立场即由此原则推导而来。

供应商提供的文档应被视为一种需要核实的主张,而绝不能视为核实本身。它在采购记录中的作用是确立供应商所声称的内容,这是问责问题,而非证明问题

独立评估应当在具有实际部署代表性的条件下进行,包括系统在实际服务中将遇到的用户群体、语言语域、主题内容以及环境条件。在其他条件下进行的评估,回答的是机构并未提出的问题。

失败情形应当在授标之前得到主动披露,并在合同中予以处理。一个失败模式未被记录在案的系统,无论附带多少文档,都不能算作已经过评估。

聋人和听障利益相关方应当参与验收标准和评估设计的制定,而不仅仅是在部署后提供反馈。在机构信心已经形成之后才让其参与,只是名义上的咨询。

失败责任应当在采购时点明确分配,并为受影响个人——而非仅仅是签约机构——提供明确的补救和升级处理机制。

实施考量

采取证据性立场的机构应当预见到三种实际的阻力。第一是时间进度。独立评估会延长采购周期,这一延长必须提前向领导层说明理由,而不能由采购部门默默承受。最有说服力的辩护方式是比较性的:将部署前评估的成本,与部署后发现故障所导致的补救、合同终止或诉讼成本相比较。

第二是能力问题。目前大多数机构并不具备为手语人工智能设计具有代表性评估条件的内部专业能力,而这种能力的缺失常常被误读为依赖供应商文档的理由。更准确的解读应当是:这正说明了对外部评估能力的需求。

The third is market response. Vendors accustomed to artifact based procurement may resist evaluation requirements, and some will decline to participate. Institutions should recognize that a vendor's unwillingness to be evaluated under deployment representative conditions is itself evidence, and is among the more informative signals available during acquisition.

实施不必一步到位才有意义。机构可以从在现有招标文件中加入证据性问题开始,将失败情形披露作为响应性标准之一,或将授标与成功通过独立评估挂钩。每一步都使机构记录从“信心”向“证据”靠近。

结语立场

人工智能治理的未来,不会由哪个组织收集了最多的合规文件来决定,而将由哪些组织建立起能够区分信心与证据的采购体系来决定。

> **第五参数**
>
> 有效治理始于程序性信心终止、独立证据开始之处。

参考文献

- Salamon, L. M. (Ed.). (2002). *The Tools of Government: A Guide to the New Governance*. Oxford University Press.
- Stone, D. (2012). *Policy Paradox: The Art of Political Decision Making* (3rd ed.). W. W. Norton.

- National Institute of Standards and Technology. (2023). *Artificial Intelligence Risk Management Framework* (AI RMF 1.0). NIST AI 100-1.
- International Organization for Standardization. (2023). ISO/IEC 42001:2023, Information technology - Artificial intelligence - Management system.
- Information Technology Industry Council. *Voluntary Product Accessibility Template*(VPAT)。依据《康复法案》第508条所引用的无障碍合规报告工具。
- 《欧洲议会和理事会关于人工智能的统一规则条例》(EU) 2024/1689。
- Novara Consulting Group. *The Fifth Parameter*, Issue 002, "The Procurement Trap."
- Novara Consulting Group. *The Fifth Parameter*, Issue 011 / Policy No. 2, "The Accountability Gap."

作者说明

Heather M. Grizzle-Odland 是 Novara Consulting Group LLC 的创始人兼首席执行官,专注于人工智能无障碍治理、手语人工智能采购、机构问责制以及公共部门技术监督。她是 Sign Language Access Trust Framework™ 的开发者,也是《The Fifth Parameter》的作者,该刊物探讨机构在评估新兴技术时常常忽视的治理问题。

她的工作聚焦于证据、问责、由聋人主导的治理,以及技术合规与现实世界无障碍之间的差距。

> **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.

引用方式

Grizzle, H. M. (2026, July 22). 采购剧场问题:程序性保障、证据性保障与无障碍AI的治理. Novara Consulting Group. https://doi.org/10.5281/zenodo.23003831