NOVARA
CONSULTING GROUP

This document was translated from English by machine and may contain mistakes. The English original is at https://www.novaracg.com/2026/07/14/who-audits-the-audit-trail-ai-generated-conformance-claims-and-the-procurement-problem/.

谁来审计审计记录?AI生成的合规声明与采购难题

Heather Grizzle · July 14, 2026

Generative AI can now write a VPAT. Why Accessibility Conformance Reports must be backed by verifiable evidence, not documents alone.

目录

1. 四份报告，一个证据空间
2. 从符合性到溯源性
3. 一种结构性失衡，而非技术性失衡
4. 证据治理
5. 作者说明

公共采购的历史，在某种程度上可以被解读为文件信任的历史。政府很少在审查源代码之后才采购软件。大学不会独立检查学习管理系统所声称的每一项安全控制措施，医院也不会例行重新进行医疗软件供应商已完成的无障碍测试。采购依赖的是介于采购方与产品本身之间的文件陈述。认证、技术报告、审计结果、符合性评估以及无障碍符合性报告之所以存在，正是因为直接验证成本高昂、专业性强，且往往不切实际。因此，机构购买的不仅是产品，还有为这些产品提供支持的证据。

这种安排一直都需要判断力。VPAT从来都不是为了证明无障碍性确实存在而设计的。它被设计为供应商依据公认无障碍标准进行评估的结构化呈现。采购官员即便高度依赖该文件本身，也理解这一区别。该报告之所以具有分量，并非因为它拥有独立的法律权威，而是因为编制该报告通常需要有人投入大量时间对产品进行评估。制作该文件所投入的努力，尽管并不完美，却成为了制作背后证据所投入努力的一种代替指标。

生成式人工智能在不改变文件外观的情况下，破坏了这种关系。

一个机构现在能够在极短的时间内，制作出技术上流畅、内部逻辑一致、格式专业且具有说服力的无障碍文件。标准被正确引用，技术术语被恰当使用，其结构与采购官员在过去二十年间审阅过的数千份符合性报告如出一辙。然而，这些特征都无法揭示该文件究竟反映的是广泛的无障碍测试、有限的内部审查，还是根本没有任何有意义的评估。写作的精巧程度，正日益与证据的可靠程度相脱节。

写作的精巧程度，正日益与证据的可靠程度相脱节。

这并非反对人工智能的论点。文件编制历来就是一种翻译行为。工程师将软件行为转化为技术性发现，无障碍专家将这些发现转化为采购语言，采购官员再将技术语言转化为采购决策。生成式人工智能只是加速了这一翻译过程中的某一环节。困难出现在翻译被误认为是验证之时。一份写得很好的报告可能准确描述了一项严谨的评估，但也可能描述的是从未经独立审查的假设。文件本身并不能提供可靠的方式来区分这两种可能性。

四份报告，一个证据空间

其后果比乍看之下更为深远。采购文件中现在包含的文件，往往在外观上大体相似，尽管它们所代表的证据基础可能截然不同。一份无障碍符合性报告可能总结了由经验丰富的无障碍专业人员采用自动化与人工相结合的方法所进行的独立测试；另一份可能只代表从未经过外部审查的内部工程评估；第三份可能准确记录了广泛的测试，同时借助生成式人工智能来提升条理与可读性；第四份则可能几乎完全根据产品规格说明和营销材料生成，而底层软件从未真正依据无障碍标准接受评估。一旦被简化为一份格式专业的报告，这些文件尽管基础迥异，却在采购记录中占据着同样的证据空间。

从符合性到溯源性

因此，传统的采购问题已不如从前那样有用。多年来，无障碍评估的核心在于产品是否符合技术标准。然而，越来越重要的问题在于该结论所依据的基础是什么。当制作有说服力文件的成本已大幅下降时，符合性声明的存在本身已不再具有多大的说明意义。采购官员反而需要了解该声明本身的来源信息：由谁进行了评估、采用了哪些方法、自动扫描是否辅以人工测试、辅助技术用户是否参与其中、结论是否可被独立复现，以及支持性证据是否仍可供外部审查。这些问题关乎证据而非无障碍本身，但它们却日益决定着无障碍声明是否值得机构信赖。

近期的采购指南已开始反映这一更广泛的转变。公共部门指南不再将文件提交视为尽职调查的终点，而是日益强调实证性、可复现性、方法透明度以及独立验证。因此，一份无障碍符合性报告的意义，与其说在于其存在本身，不如说在于它能否引导采购官员找到经得起有意义审查的证据。该报告之所以有价值，并非因为它宣称合规，而是因为它解释了为何这一宣称应当被采信。

一种结构性失衡，而非技术性失衡

其影响远远超出无障碍采购的范畴。人工智能已大幅降低了在几乎所有治理领域中制作有说服力的机构文件的成本。安全评估、风险分析、治理声明、实施计划、隐私文件以及采购叙述，都可以以日益流畅的方式、以日益减少的努力起草完成。然而，这些文件所依赖的实际工作，却依旧难以获得同等程度的加速。独立测试仍需合格的评估人员，验证仍依赖可复现的方法论，证据依然源自观察而非撰写。由此产生的失衡因而是结构性的，而非技术性的。陈述的生产速度，已远远超过了这些陈述所描述之事实的生产速度。

证据治理

正是在这样的背景下，证据治理变得日益重要。手语无障碍信任框架（Sign Language Access Trust，SLAT Framework）的开发前提是：文件不应脱离产生它的机构流程而被独立评估。该框架并不将供应商的声明视为自证成立，而是审查治理结构、验证方法、透明度做法、披露质量，以及支持性证据的来源。核心问题不在于是否作出了某项声明，而在于是否存在足够的证据，使一个理性的采购方能够在做出影响无障碍性、公共支出及法律问责等重大决策时，合理依赖该声明。

> **真正重要的检验标准**
>
> 核心问题不在于是否作出了某项声明，而在于是否存在足够的证据，使一个理性的采购方能够依赖该声明。

公共采购长期以来依赖文件性证物，因为机构无法独立验证呈交给它们的每一项技术性主张。这种依赖不太可能消失。发生改变的，是文件与制作该文件历来所需投入的努力之间的关系。随着人工智能持续降低具有说服力的技术写作的成本，采购实践必将更加重视支撑这些文件的证据的质量、可追溯性与独立性。无障碍采购未来的可信度，将不

再主要取决于符合性报告的精巧程度，而是取决于各机构区分文件陈述与可证实事实的能力。

作者说明

关于作者

Heather M. Grizzle-Odland

Heather M. Grizzle-Odland 是 Novara Consulting Group 的创始人兼首席执行官。她的工作聚焦于人工智能无障碍治理、手语人工智能采购风险、证据标准以及由Deaf群体主导的问责机制。她还是手语无障碍信任框架（Sign Language Access Trust，SLAT Framework）的开发者，这是一套用于评估手语人工智能系统治理、验证、透明度及采购就绪程度的结构化评估模型。

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 14). 谁来审计审计记录?AI生成的合规声明与采购难题. Novara Consulting Group. https://doi.org/10.5281/zenodo.23003841