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/.

Das Problem des Beschaffungstheaters: Verfahrenssicherheit, evidenzbasierte Sicherheit und die Governance von Barrierefreiheits-KI

Heather Grizzle · July 22, 2026

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

Inhalt

1. Zusammenfassung
2. Das politische Problem
3. Institutionelle Analyse
4. Governance-Implikationen
5. NCG-Grundsatzposition
6. Umsetzungsüberlegungen
7. Abschließende Position
8. Quellen
9. Anmerkung der Autorin

Zusammenfassung

Institutionen, die Systeme künstlicher Intelligenz für den Kommunikationszugang erwerben, sind zunehmend überzeugt, das Risiko im Griff zu haben – und diese Überzeugung ist zunehmend unbegründet. Sie stützt sich auf Artefakte: ausgefüllte Checklisten, Herstellerdokumentationen, einstudierte Vorführungen, ausgehandelte Gewährleistungsklauseln, eingereichte Barrierefreiheitserklärungen. Jedes Artefakt ist

echt. Keines davon, weder einzeln noch zusammengenommen, ist darauf ausgelegt, die Frage zu beantworten, die die Institution eigentlich beantwortet haben muss: ob das System im Einsatzkontext, für den es beschafft wird, sicher, zuverlässig und angemessen funktionieren wird.

Diese Ausgabe bezeichnet diesen Zustand als Beschaffungstheater und argumentiert, dass er struktureller und nicht böswilliger Natur ist. Beschaffungsfachleute werden für verfahrenstechnische Vollständigkeit belohnt. Anbieter werden für überzeugende Vorführungen belohnt. Beide Seiten können ihre Pflichten vollständig erfüllen, während die zentrale Leistungsfrage ungeprüft bleibt. Die Abhilfe besteht nicht in mehr Dokumentation. Sie besteht in einer Verschiebung des maßgeblichen Standards von der verfahrensbasierten Absicherung, die fragt, ob die vorgeschriebenen Schritte durchgeführt wurden, hin zur evidenzbasierten Absicherung, die fragt, ob die gesammelten Belege in der Lage waren, die Wahrheit aufzudecken.

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.

Das politische Problem

Governance wird nicht an Aktivität gemessen. Sie wird daran gemessen, ob die durchgeführten Aktivitäten in der Lage sind, die Wahrheit aufzudecken. Die meisten aktuellen Beschaffungsprozesse für barrierefreie KI scheitern an diesem Test – nicht weil der Prozess nachlässig wäre, sondern weil er für eine andere Produktklasse konzipiert wurde.

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.



Drei Vermengungen halten das Problem am Leben. Dokumentation wird als Validierung behandelt, obwohl Dokumentation festhält, was ein Anbieter behauptet, nicht was ein System tatsächlich leistet. Zertifizierung wird als Leistung behandelt, obwohl Zertifizierung typischerweise die Verfahrenskonformität bescheinigt, nicht die Qualität der Ausgabe. Konformität wird als Wirksamkeit behandelt, obwohl Konformität die Übereinstimmung mit einem technischen Standard misst, nicht ob eine gehörlose oder schwerhörige Person tatsächlich funktionalen kommunikativen Zugang erlangt hat. Ein Beschaffungsprozess, der diese drei Konzepte nicht unterscheiden kann, erzeugt institutionelles Risiko im direkten Verhältnis zur Anzahl der ausgefüllten Formulare, denn jedes ausgefüllte Formular erhöht das Vertrauen, ohne das Wissen zu vermehren.

Institutionelle Analyse

Die Anreizstruktur erklärt das Fortbestehen dieses Musters überzeugender als jede Erklärung durch böse Absicht. Beschaffungsstellen operieren unter Prüfregimen, die untersuchen, ob die vorgeschriebenen Schritte stattgefunden haben, nicht ob diese Schritte erkenntnistheoretisch angemessen waren. Eine Akte mit einem Konformitätsbericht des Anbieters zur Barrierefreiheit, einem Vorführungsprotokoll und ausgehandelten Vertragsklauseln übersteht eine Prüfung. Eine Akte, die dokumentiert, dass die Institution das System eigenständig an repräsentativen gehörlosen Nutzerinnen und Nutzern unter repräsentativen Bedingungen getestet und für unzureichend befunden hat, erzeugt Zeitplanrisiken und politische Angriffsfläche. Das System belohnt die erste Akte.

Anbieter stehen vor einem spiegelbildlichen Anreiz. Märkte für barrierefreie KI belohnen Vorführungsqualität, narrative Kohärenz und Geschwindigkeit bis zum Vertragsabschluss. Anbieter werden selten aufgefordert, Versagensbedingungen offenzulegen, und eine einseitige Offenlegung verursacht wettbewerbliche Kosten. Öffentlich zugängliche Belege stützen derzeit nicht die Behauptung, dass gegenwärtige kommerzielle Gebärdensprach-KI-Systeme gegen unabhängig konstruierte, einsatzrepräsentative Testbedingungen mit veröffentlichten Ergebnissen evaluiert wurden. In Ermangelung solcher Belege werden Marketingaussagen und selbst erstellte Dokumentation standardmäßig zur institutionellen Aktenlage.

Die betroffene Bevölkerungsgruppe trägt das verbleibende Risiko. Gehörlose und schwerhörige Nutzerinnen und Nutzer sind in der Regel keine Beteiligten am Beschaf-

fungsprozess, werden bei der Abnahmeprüfung selten einbezogen und begegnen dem System erst, nachdem das institutionelle Vertrauen bereits hergestellt und der Vertrag geschlossen wurde. Zu diesem Zeitpunkt hat sich die Beweislast umgekehrt. Der Nutzer muss nachweisen, dass das System versagt hat, gegen eine institutionelle Aktenlage, die behauptet, es sei Sorgfalt walten gelassen und das System als konform befunden worden. Diese Umkehrung ist der eigentliche Schaden des Beschaffungstheaters, und sie wird vollständig von legitimen Akteuren erzeugt, die legitime Verfahren befolgen.

Im Begriffsrahmen der Regierungslehre betrachtet, ist Beschaffung keine administrative Formalität, sondern ein Instrument öffentlichen Handelns, und die Wahl des Instruments bestimmt, wer befugt ist zu sprechen, was als Beweis gilt und wo Rechenschaftspflicht letztlich verortet wird. Liest man dies zusammen mit der Literatur zu kausalen Erzählungen, so ist die vorherrschende Darstellung des Versagens barrierefreier KI eine Unfallgeschichte, in der unberechenbare Technologie gelegentlich versagt und kein Akteur die Verantwortung trägt. Die hier vorgebrachte evidenzbasierte Rahmung stützt eine andere Darstellung, mindestens eine der Fahrlässigkeit: Die Versagensfälle sind vorhersehbar, die Bedingungen, unter denen sie auftreten, sind im Voraus erkennbar, und die Entscheidung, nicht hinzusehen, ist selbst eine Governance-Entscheidung.

Die Entscheidung, nicht hinzusehen, ist selbst eine Governance-Entscheidung.

Governance-Implikationen

Der Übergang von verfahrensbasierter zu evidenzbasierter Absicherung erfordert von Institutionen, einen anderen Fragenkatalog zu stellen und die Antworten als Abnahmekriterien zu behandeln, nicht als Hintergrundinformation. Welche Belege stützen die gemachte Behauptung, und handelt es sich dabei um Leistungsbelege oder um Verfahrensbelege? Wer hat die Belege erzeugt, in welchem kommerziellen Verhältnis zum Ergebnis, und mit welcher methodischen Transparenz? Kann das Ergebnis von einer Partei ohne Interesse am Erfolg der Beschaffung unabhängig reproduziert werden? Unter welchen konkreten Bedingungen versagt das System, mit welcher Rate, und mit welcher Schweregradverteilung über verschiedene Nutzergruppen hinweg?

Und wenn diese Versagensfälle im Einsatz auftreten, welche Partei trägt die Verantwortung, durch welchen Mechanismus, in welchem Zeitrahmen, und welche Abhilfe steht der betroffenen Person zur Verfügung?

Diese Fragen heben nicht nur die Beweisanforderungen an. Sie verlagern die Rechenschaftspflicht. Eine Beschaffungsakte, die sie beantwortet, legt fest, was die Institution wusste, wann sie es wusste und was sie akzeptierte. Eine Beschaffungsakte, die sie nicht beantwortet, belegt lediglich, dass die Institution ein Verfahren abgeschlossen hat. Angesichts aufkommender Governance-Erwartungen, einschließlich risikobasierter Rahmenwerke, die von Organisationen verlangen, den Einsatzkontext zu charakterisieren und Restrisiken zu dokumentieren, wird die zweite Akte zunehmend schwerer zu verteidigen.

Die entscheidende Kompetenz ist die institutionelle Fähigkeit, Vertrauen von Beweisen zu unterscheiden. Diese Fähigkeit wohnt nicht in Checklisten. Sie wohnt in Evaluationsinstrumenten, in der Fähigkeit, repräsentative Testbedingungen festzulegen, und in der Bereitschaft, eine Anbieterbehauptung als Hypothese und nicht als Befund zu behandeln.

NCG-Grundsatzposition

Novara Consulting Group vertritt die Auffassung, dass die Beschaffung barrierefreier KI von evidenzbasierter Absicherung geleitet werden sollte, und dass sich die folgenden Positionen aus diesem Prinzip ergeben.

Vom Anbieter erstellte Dokumentation sollte als eine Behauptung behandelt werden, die der Überprüfung bedarf, niemals als Überprüfung selbst. Ihre Funktion in der Beschaffungsakte besteht darin festzuhalten, was der Anbieter behauptet hat – eine Frage der Rechenschaftspflicht, nicht des Beweises.

Unabhängige Evaluierung sollte unter Bedingungen durchgeführt werden, die für den tatsächlichen Einsatzkontext repräsentativ sind, einschließlich der Nutzergruppen, sprachlichen Register, Themenbereiche und Umgebungsbedingungen, denen das System im Einsatz begegnen wird. Eine Evaluierung unter anderen Bedingungen beantwortet eine Frage, die die Institution nicht gestellt hat.

Versagensbedingungen sollten aktiv offengelegt und vor der Vergabe vertraglich geregelt werden. Ein System, dessen Versagensmodi nicht dokumentiert sind, wurde nicht evaluiert, unabhängig davon, wie viel Dokumentation es begleitet.

Gehörlose und schwerhörige Interessenvertreter sollten an der Festlegung der Abnahmekriterien und der Gestaltung der Evaluierung beteiligt werden, nicht nur am Feedback nach dem Einsatz. Beteiligung, nachdem sich institutionelles Vertrauen bereits gebildet hat, ist Konsultation nur dem Namen nach.

Die Verantwortung für Versagen sollte bereits beim Erwerb ausdrücklich zugewiesen werden, mit definierter Abhilfe und Eskalationsmöglichkeit für die betroffene Person, nicht nur für die auftraggebende Institution.

Umsetzungsüberlegungen

Institutionen, die eine evidenzbasierte Haltung einnehmen, sollten mit drei praktischen Reibungspunkten rechnen. Der erste ist der Zeitplan. Unabhängige Evaluierung verlängert die Beschaffungszyklen, und diese Verlängerung muss der Führungsebene vorab begründet werden, statt stillschweigend von der Beschaffungsstelle aufgefangen zu werden. Die tragfähigste Begründung ist vergleichend: die Kosten der Evaluierung vor dem Einsatz gegenüber den Kosten der Nachbesserung, Vertragsbeendigung oder von Rechtsstreitigkeiten infolge von im Einsatz entdeckten Mängeln.

Der zweite ist die Kompetenz. Die meisten Institutionen verfügen derzeit nicht über die interne Expertise, um repräsentative Evaluierungsbedingungen für Gebärdensprach-KI zu gestalten, und das Fehlen dieser Expertise wird häufig fälschlich als Grund angesehen, sich auf die Dokumentation des Anbieters zu verlassen. Zutreffender ist es, darin eine Spezifikation für externe Evaluierungskapazität zu sehen.

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.

Die Umsetzung muss nicht vollständig sein, um bedeutsam zu sein. Institutionen können beginnen, indem sie bestehenden Ausschreibungsinstrumenten evidenzbasierte Fragen hinzufügen, die Offenlegung von Versagensbedingungen als Eignungskriterium verlangen oder die Vergabe von einer erfolgreichen unabhängigen Evaluierung abhängig machen. Jeder Schritt bewegt die institutionelle Aktenlage von Vertrauen hin zu Beweisen.



Abschließende Position

Die Zukunft der KI-Governance wird nicht dadurch bestimmt, welche Organisation die größte Zahl an Compliance-Dokumenten sammelt. Sie wird dadurch bestimmt, welche Organisationen Beschaffungssysteme aufbauen, die in der Lage sind, Vertrauen von Beweisen zu unterscheiden.

> **DER FÜNFTE PARAMETER**
>
> Wirksame Governance beginnt dort, wo verfahrensbasiertes Vertrauen endet und unabhängige Beweisführung beginnt.

Quellen

- Salamon, L. M. (Hrsg.). (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* (3. Aufl.). 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). Instrument zur Berichterstattung über Konformität mit Barrierefreiheit, referenziert unter Section 508 des Rehabilitation Act.
- Verordnung (EU) 2024/1689 des Europäischen Parlaments und des Rates zur Festlegung harmonisierter Vorschriften für künstliche Intelligenz.
- 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."

Anmerkung der Autorin

Heather M. Grizzle-Odland ist Gründerin und CEO von Novara Consulting Group LLC, wo sie sich auf KI-Governance im Bereich Barrierefreiheit, die Beschaffung von Gebärdensprach-KI, institutionelle Rechenschaftspflicht und die technologische Aufsicht im öffentlichen Sektor konzentriert. Sie ist die Entwicklerin des Sign Language Access Trust Framework™ und Autorin von The Fifth Parameter, einer Publikation, die sich mit den Governance-Fragen befasst, die Institutionen bei der Bewertung neuer Technologien häufig übersehen.

Ihre Arbeit konzentriert sich auf Evidenz, Rechenschaftspflicht, von Gehörlosen geleitete Governance und die Kluft zwischen technischer Konformität und tatsächlicher Barrierefreiheit in der Praxis.

> **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.

Zitierweise

Grizzle, H. M. (2026, July 22). Das Problem des Beschaffungstheaters: Verfahrenssicherheit, evidenzbasierte Sicherheit und die Governance von Barrierefreiheits-KI. Novara Consulting Group. https://doi.org/10.5281/zenodo.23003831