

CLOSING THE TRANSPARENCY GAP

**Disability-Inclusive and
Ethical AI in Public Services**
Evidence from 34 Countries



Transparency is the minimum.
Accessible and ethical AI
is a requirement.



ETHICS.

AI that respects
human rights



ACCESSIBILITY.

AI that works
for everyone



ACCOUNTABILITY.

AI that is monitored
and auditable



HUMAN RIGHTS.

AI that protects the rights
of persons with disabilities



AI GOVERNANCE FOR AN INCLUSIVE FUTURE

No one left behind.



Closing the Transparency Gap for Disability and Ethical AI in Public Services

Report researched and prepared by: Martin Gould, Director of Research, G3ict.

Edited by Axel Leblois, G3ict Founder and President Emeritus.

Reviewed by:

- Christopher Lee, PhD, President & CEO, G3ict
- Ann Junker and Olivia Pollard, Subject Matter Experts, members of the Organizational Professional Development Committee (OPDC) of the International Association of Accessibility Professionals (IAAP), a division of G3ict.

Table of Contents

Foreword.....	4
Introduction.....	5
The Global Transparency Gap	6
34 Countries out of 193 Publish Evidence.....	6
Methodology.....	7
Evidence Collection	8
Indicators for Accessible and Ethical AI in Public Services	9
Indicator Matrix of 34 Countries Providing Evidence	11
Transparency Country Ranking Table	13
Flying Blind: The Consequences of Limited AI Transparency in Public Services	13
Call for Action for Public Services	14
Pathways for Governments	15
Pathways for Advocates.....	15
Appendices Appendix 1: G3ict Recommended AI Risk Taxonomy for Persons with Disabilities and Harm Repository	16
Risk Taxonomy	16
Harms Repository.....	20
Appendix 2: Use Case Scenarios.....	21
2.1 Telehealth AI Misinterprets AAC Speech	21
2.2: Voice Authentication Rejects Older Adults	22
2.3: Navigation App Routes Wheelchair Users Through Inaccessible Paths	24
Bibliography.....	27

Foreword

Many of the technologies the world now takes for granted, from captioning to voice control to predictive text, trace a real part of their origin to tools built with and for persons with disabilities. We have every reason to welcome artificial intelligence in public services. We also have every reason to ask who is checking whether it works for us. This report answers that question, and the answer is sobering. Of 193 UN Member States, only 34 publish enough evidence to show how their public-sector AI is governed. The remaining 159 may be doing careful work or none at all. From the outside, no one can tell. For the 1.3 billion persons with disabilities worldwide, that silence is itself a risk. An eligibility algorithm or identity check that cannot be examined cannot be challenged.

One finding stands out for IDA and our members. Among the 34 countries with published evidence, only five document the participation of persons with disabilities in AI governance. Article 4.3 of the Convention on the Rights of Persons with Disabilities requires States to closely consult and actively involve us, through our representative organizations, in the decisions that affect our lives. Few decisions will shape our lives more in the coming decade than how governments design, buy and oversee AI.

A commitment on paper is a structure. What happens after, in procurement, oversight and practice, is what turns it into something real. Much of the AI in public services is built by private companies and bought by governments, which makes procurement the place to write accessibility in from the start. For industry, our participation is an asset: the people most exposed to automated error are often the first to find it.

I urge governments to treat these findings as a baseline and to publish more before the next edition. And I urge our movement to use them: to demand the evidence, to take our place where these systems are designed, and to make our presence there as indispensable as the technologies we helped invent.

José Viera

Executive Director, International Disability Alliance

Introduction

AI is remodeling public services and civic life, creating significant opportunities for government innovation across sectors, including Smart Cities, while supporting more accessible user interfaces and assistive technologies.

Yet as public agencies expand their use of AI, few governments publish verifiable evidence that these systems are accessible, ethically governed, protect the rights of persons with disabilities, and are subject to reasonable oversight and monitoring. Critical errors such as telehealth AI misinterpreting AAC speech, voice authentication rejecting older adults, or navigation apps routing wheelchair users through inaccessible paths must be identified before systems are deployed.

The Convention on the Rights of Persons with Disabilities (CRPD) affirms that disability rights are human rights, and Article 9 explicitly requires accessible information and communication technologies. The UN Global Digital Compact similarly embeds principles for ethical and accessible AI. Yet governments worldwide continue to develop AI systems with few or no documented safeguards, leaving millions of persons with disabilities at risk of being unable to access, challenge, or trust essential public services—and, in the worst cases, of being inadvertently harmed by them. These issues are especially urgent given that 1.3 billion people worldwide live with disabilities - a number that is rising rapidly as populations age.

The first section of this report documents the global transparency gap among 193 UN member countries and of the consequences of limited AI transparency in public services. Following sections include a Call for Action for Public Services and suggested pathways for governments and advocates. Appendix One presents G3ict's recommended taxonomy of risks and Appendix 2 three detailed use cases and their diagnostic.

The Global Transparency Gap

Of the 193 UN Member States, as of August 2026, only 34 publish enough evidence to demonstrate even a basic level of transparency in AI governance for public services.

These 34 countries are the only ones that make available:

- Public AI strategies.
- AI regulations or risk frameworks.
- Accessibility or disability-rights laws.
- Digital public infrastructure rules.
- Evidence of audits, oversight, or outcomes.

The remaining 159 countries provide no verifiable evidence of implementing ethical and accessible AI in public services.

G3ict Research believes that calling for greater data transparency, while benchmarking countries' progress in providing evidence, will draw governments' attention to this gap and encourage systematic, structured, and publicly available reviews of AI applications in public services to ensure ethical and accessibility compliance needed for remediation.

34 Countries out of 193 Publish Evidence

These are the only nations transparent enough to be evaluated on AI accessibility governance.

Europe (21 countries): Austria, Belgium, Czechia, Denmark, Estonia, Finland, France, Germany, Iceland, Ireland, Italy, Luxembourg, Netherlands, Norway, Poland, Portugal, Slovenia, Spain, Sweden, Switzerland, United Kingdom.

Anglophone OECD (4 countries): Australia, Canada, New Zealand, and the United States.

Asia (4 countries): Israel, Japan, Korea (Republic of), Singapore.

Middle East (1 country): United Arab Emirates.

Latin America (3 countries): Brazil, Chile, Mexico.

Africa (1 country): Rwanda.

Other countries not listed above are developing AI applications for public services without sufficient transparency to evaluate them.

82% of countries, representing 159 nations, provide no evidence of ethical and accessible AI implementation in public services.

Methodology

The G3ict Structure-Process-Outcomes (S-P-O) indicators in this report are based exclusively on publicly available and verifiable government evidence. Analysts assess what governments publicly report about their AI governance, accessibility, and oversight practices using a standardized desk-research methodology. The analysis does not rely on internal government documents, unpublished evidence, or automated data scraping. Scores are derived from original research conducted for this report using official government sources, rather than external datasets or third-party assessments.

For Structure indicators, a country scores well only if the relevant documents are publicly accessible and demonstrate explicit protections for persons with disabilities. For Process indicators, most governments do not publish evidence of enforcing accessibility in AI systems; so, even if they do, Process scores tend to be lower. For Outcomes, fewer countries publish evidence, making this the weakest of the three groups of indicators of the SPO framework.

Evidence Collection

Analysts examine:

- Official government websites.
- Legal repositories.
- Procurement portals.
- Digital policy platforms.
- Public audit reports

Alternative approaches such as surveys, interviews, and expert assessments were considered. However, a public-evidence review was selected because it provides the most consistent and replicable method for comparing all 193 UN Member States using verifiable information published by governments.

Only publicly available, accessible, and verifiable evidence is considered. Governments that do not publish such evidence receive lower scores, regardless of any internal activities. Consequently, only 34 countries qualify for evaluation, as they provide sufficient evidence for assessment. Evidence did not need to appear on a government's primary website. When official sources referenced or linked to relevant documents located in other official repositories or agencies, those materials were reviewed and considered as part of the assessment.

All 193 UN Member States are evaluated using a standardized, transparency-based desk-research methodology. Evidence is collected exclusively from publicly available government sources, such as national AI strategies, digital-government portals, legislative repositories, procurement regulations, accessibility standards, and official policy documents published in each country's primary languages.

Native-language search queries and machine-translation tools identify and review non-English materials, ensuring comprehensive coverage across diverse legal and administrative systems. All findings are scored using binary criteria aligned with the Structure–Process–Outcomes (S–P–O) framework. Structure indicators capture the

existence of laws, strategies, and institutional mandates. Process indicators assess implementation mechanisms, including procurement rules, disclosure requirements, and oversight procedures. Outcome indicators verify whether governments publish results, audits, or accessibility-related AI documentation. Each data point undergoes internal assessment to secure accuracy, consistency, and replicability. This transparency-based approach is consistent with methodologies used by other global AI governance indexes, such as the Global Public Sector AI Index (GPSAI), the CAIDP AI and Democratic Values Index, and the UN DESA Digital Government assessments, which also rely on open-source, multilingual government evidence to produce globally comparable benchmarks of public-sector AI readiness and accountability.

Indicators for Accessible and Ethical AI in Public Services

Structure Domain: Governance Foundations (Maximum 5 points)

- S1. The national AI strategy contains explicit commitments to accessibility.
- S2. The AI regulatory framework entails obligations aligned with the Convention on the Rights of Persons with Disabilities (CRPD).
- S3. The AI risk-classification system identifies disability-related impacts, including impacts affecting older persons living with disabilities.
- S4. Digital Public Infrastructure (DPI) governance includes explicit accessibility requirements.
- S5. Public procurement regulations require AI systems to be accessible.

Process Domain: Implementation Mechanisms (Maximum 6 points)

- P1. Accessibility reviews are systematically conducted for public-sector AI deployments.
- P2. Dataset audits incorporate checks for disability representation.
- P3. Protocols for testing multimodal interactions are documented and implemented.
- P4. Mechanisms for human monitoring are established to govern automated decision-making.

- P5. Pathways for redress are available for algorithm-based harm affecting persons with disabilities.
- P6. Persons with disabilities participate in formal AI governance processes.

Outcomes Domain: Published Accessibility and Equity Evaluations (Maximum 4 points)

- O1. Evaluations of public services using AI, measured against documented accessibility standards, are publicly available and accessible.
- O2. Accessibility compliance tests for digital identity and authentication systems are publicly available and accessible.
- O3. Evaluations of accessibility requirements for AI-powered interfaces are publicly available and accessible.
- O4. Documented reviews of algorithmic equity outcomes for persons with disabilities are publicly available and accessible.

The Structure-Process-Outcomes (S-P-O) indicators are derived from a manual review of publicly available government evidence. This model assesses whether countries publish laws, enforcement systems, and data regarding real-world accessibility and ethical AI outcomes.

The countries' scores are based on primary government evidence. International frameworks, AI governance research, accessibility standards, and related studies informed the development of the indicators and the recommended risk taxonomy and are listed in the bibliography.

The S-P-O Methodological Framework

The Office of the High Commissioner for Human Rights (OHCHR) developed the Structure–Process–Outcome (S-P-O) conceptual and methodological framework, which can be applied and adapted at the national level. This framework provides a standardized approach for identifying indicators to monitor civil and political rights as well as economic, social, and cultural rights.

The S-P-O framework was selected because it distinguishes between government commitments, implementation mechanisms, and documented results. This allows policymakers to identify where accessibility and accountability gaps occur and what actions may be required to address them.

The framework recommends developing structural, process, and outcome indicators. This configuration of indicators helps assess the steps being taken by States in addressing their obligations – from commitments and acceptance of international human rights standards (structural indicators) to efforts being made to meet the obligations that flow from the standards (process indicators) and on to the results of those efforts (outcome indicators).

G3ict Research aligns its global analysis, whenever possible, with the S-P-O methodological framework, including for its Digital Accessibility Rights Evaluation Index (DARE Index) and for this report on transparency gaps among countries to ensure ethical and accessible AI in public services.

Indicator Matrix of 34 Countries Providing Evidence

The DARE Index framework examines the progression from commitments to implementation and outcomes. This report adapts that approach to AI governance by assessing whether governments publish evidence of accessibility commitments, implementation measures, and documented results.

All scores are calculated using the DARE sequential hierarchy model and validated against aggregate source totals. In the table, ● indicates presence, and ○ indicates absence

Country	Structure						Process							Outcomes					Σ
	S1	S2	S3	S4	S5		P1	P2	P3	P4	P5	P6		O1	O2	O3	O4		
Australia	●	●	●	●	○	4	●	●	●	●	○	○	4	●	●	●	●	4	12
Austria	●	●	●	●	●	5	●	●	●	●	●	○	5	●	●	●	○	3	13
Belgium	●	●	●	●	●	5	●	●	●	●	●	○	5	●	●	●	○	3	13
Brazil	●	●	●	●	○	4	●	●	●	○	○	○	3	●	●	○	○	2	9
Canada	●	●	●	●	○	4	●	●	●	●	○	○	4	●	●	●	●	4	12
Chile	●	●	●	●	○	4	●	●	●	●	○	○	4	●	●	○	○	2	10
Czechia	●	●	●	●	●	5	●	●	●	●	●	○	5	●	●	●	○	3	13
Denmark	●	●	●	●	●	5	●	●	●	●	●	●	6	●	●	●	●	4	15
Estonia	●	●	●	●	●	5	●	●	●	●	●	●	6	●	●	●	●	4	15
Finland	●	●	●	●	●	5	●	●	●	●	●	○	5	●	●	●	○	3	13
France	●	●	●	●	●	5	●	●	●	●	●	○	5	●	●	●	○	3	13
Germany	●	●	●	●	●	5	●	●	●	●	●	○	5	●	●	●	○	3	13
Iceland	●	●	●	●	●	5	●	●	●	●	●	○	5	●	●	●	○	3	13
Ireland	●	●	●	●	●	5	●	●	●	●	●	○	5	●	●	●	○	3	13
Israel	●	●	●	●	○	4	●	●	●	●	○	○	4	●	●	●	●	4	12
Italy	●	●	●	●	●	5	●	●	●	●	●	○	5	●	●	●	○	3	13
Japan	●	●	●	●	○	4	●	●	●	●	○	○	4	●	●	●	○	3	11
Korea (Rep. of)	●	●	●	●	○	4	●	●	●	●	○	○	4	●	●	●	○	3	11
Luxembourg	●	●	●	●	●	5	●	●	●	●	●	○	5	●	●	●	○	3	13
Mexico	●	●	●	○	○	3	●	●	●	○	○	○	3	●	●	○	○	2	8
Netherlands	●	●	●	●	●	5	●	●	●	●	●	●	6	●	●	●	●	4	15
New Zealand	●	●	●	●	○	4	●	●	●	●	○	○	4	●	●	●	●	4	12
Norway	●	●	●	●	●	5	●	●	●	●	●	●	6	●	●	●	○	3	14
Poland	●	●	●	●	○	4	●	●	●	○	○	○	3	●	●	●	○	3	10
Portugal	●	●	●	●	●	5	●	●	●	●	●	○	5	●	●	●	○	3	13
Rwanda	●	●	●	○	○	3	●	●	●	○	○	○	3	●	○	○	○	1	7
Singapore	●	●	●	●	●	5	●	●	●	●	●	○	5	●	●	●	○	3	13
Slovenia	●	●	●	●	●	5	●	●	●	●	●	○	5	●	●	●	○	3	13
Spain	●	●	●	●	●	5	●	●	●	●	●	○	5	●	●	●	○	3	13
Sweden	●	●	●	●	●	5	●	●	●	●	●	●	6	●	●	●	●	4	15
Switzerland	●	●	●	●	●	5	●	●	●	●	●	○	5	●	●	●	○	3	13
UAE	●	●	●	●	○	4	●	●	●	●	○	○	4	●	●	○	○	2	10
United Kingdom	●	●	●	●	●	5	●	●	●	●	●	○	5	●	●	●	○	3	13
United States	●	●	○	○	○	2	●	●	○	○	○	○	2	●	○	○	○	1	5
COLUMN TOTALS (●)	34	34	32	30	22	—	34	34	34	30	22	5	—	34	32	28	14	—	—

Transparency Country Ranking Table

Four countries achieved the highest score of 15: Denmark, Estonia, the Netherlands, and Sweden. Norway scored 14. Sixteen countries scored 13: Austria, Belgium, Czechia, Finland, France, Germany, Iceland, Ireland, Italy, Luxembourg, Portugal, Singapore, Slovenia, Spain, Switzerland, and the United Kingdom. The remaining countries scored below 13: Australia, Canada, Israel, and New Zealand (12); Japan and Korea (11); Chile, Poland, and the United Arab Emirates (10); Brazil (9); Mexico (8); Rwanda (7); and the United States (5). Inclusion in the assessment was based on the availability of public evidence, not on achieving a specific score threshold.

Among the countries scoring below 13, the most common gaps were the absence of publicly documented participation by persons with disabilities in AI governance processes (P6), limited evidence of accessibility requirements in AI procurement (S5), and limited publication of outcome-level evaluations, particularly algorithmic equity reviews (O4). The column totals suggest that governance foundations are more commonly documented than implementation and outcome measures, indicating that many countries have established policy commitments but publish less evidence of enforcement, monitoring, and measurable results.

Flying Blind: The Consequences of Limited AI Transparency in Public Services

Only 34 countries currently publish information on the governance, processes, and outcomes of their AI systems at a level sufficient to support at least a minimum transparency assessment for public services.

Transparency alone does not demonstrate that disability rights are being upheld. However, public reporting, audits, and evaluations are necessary to identify shortcomings, measure progress, and support corrective action when harms occur.

Among these countries, publicly available data still show significant implementation gaps. Nevertheless, measuring, documenting, and publishing outcomes gives their governments a foundation to identify shortcomings, strengthen oversight, and implement corrective measures.

In countries lacking such monitoring and reporting, the rapid deployment of artificial intelligence in public services outpaces government officials' and developers' capacity to ensure these systems uphold the rights of persons with disabilities. Without systematic and transparent accountability mechanisms, only limited evidence shows that accessibility and disability rights are effectively integrated into the design, procurement, deployment, and oversight of artificial intelligence systems. Appendices 1, 2, and 3 illustrate the consequences of these gaps.

Call for Action for Public Services

“Lack of focus on disability-inclusive, ethical, and accessible AI is not just a technical problem. It is a human rights problem.”

Christopher Lee, PhD, CEO & President, G3ict

AI systems decide who can access digital identity, who can apply for jobs, who gets public benefits, who can use transportation, and who can participate in civic life. If these systems are inaccessible or unethical, persons with disabilities are excluded from the future and may be harmed.

AI without accessibility is discrimination. AI without ethical guardrails may be harmful. AI without transparency is dangerous.

Transparency is the Minimum.
Accessible and Ethical AI is the Requirement.

Pathways for Governments

We call on every government to meet these three essential demands without exception.

Publish

- Release all AI governance documents, including strategies, regulations, risk frameworks, and DPI rules.
- Publish accessibility commitments, including CRPD implementation, digital accessibility standards, and procurement rules.
- Publish data on outcomes, including audits, oversight mechanisms, redress pathways, and accessibility results

Protect

- Enforce the CRPD within all digital environments.
- Mandate accessibility requirements in all AI procurement processes.
- Establish clear redress pathways for persons with disabilities.

Participate

- Ensure the inclusion of persons with disabilities in all AI governance processes.
- Provide funding for disability-led AI accountability initiatives.
- Require the co-design of AI-powered public services with persons with disabilities.

Pathways for Advocates

Three demands. To every government. No exceptions.

Publish

- Release all AI governance strategies.
- Publish accessibility commitments.
- Disclose audit results and outcomes.

Protect

- Enforce CRPD in digital environments.
- Mandate accessibility in AI procurement.
- Create redress pathways for disabled people.

Participate

- Include persons with disabilities in AI governance.
- Fund disability-led AI accountability work.
- Require co-design of AI public services.

AI Is Global. Accessibility and Ethics Must Be Global Too.

Appendices

Appendix 1: G3ict Recommended AI Risk Taxonomy for Persons with Disabilities and Harm Repository

Risk Taxonomy

The **G3ict Recommended Disability-Inclusive AI Risk Taxonomy** supports auditing and evaluating AI systems, including those implemented in public services. This taxonomy offers a structured approach to identifying and assessing risks that disproportionately impact persons with disabilities and older persons.

The taxonomy defines key risk categories, outlines their application, and introduces a coding system to facilitate traceability across audits, scenario testing, and validation processes. It provides a standardized framework, aligned with internationally recognized principles for accessibility, human rights, and responsible AI, for identifying, assessing, and mitigating disability-related AI risks. The taxonomy also supports consistent risk management throughout the AI lifecycle, including system updates, monitoring, and retirement decisions.

It functions as a standardized, globally aligned framework for identifying, assessing, and mitigating disability-related AI risks.

The taxonomy supports the following objectives:

- Consistency across audits, benchmarks, and scenario tests
- Traceability across evidence sources and governance decisions
- Comparability across systems, sectors, and countries

The taxonomy aligns with major international frameworks, including:

- Convention on the Rights of Persons with Disabilities (CRPD)
- Global Digital Compact (GDC)
- OECD AI Principles
- UNESCO Recommendations on the Ethics of AI
- National AI governance frameworks

Structure of the Taxonomy

The taxonomy is organized into eight primary risk categories. Each category contains subcategories, definitions, examples, and standardized risk codes. Collectively, these elements address the full lifecycle of AI systems and incorporate the lived experiences of persons with disabilities and older persons.

Risk codes should be reviewed periodically to reflect evolving technologies, regulations, and evidence. New risks may be added, existing risks may be revised, and outdated risks may be retired while preserving historical traceability.

1. Exclusion and Accessibility Failures

Risks where AI systems exclude users, misinterpret inputs, or fail to provide necessary accommodations.

- Missing accessibility features
- Inaccessible inputs or interfaces
- Inaccurate navigation systems
- Exclusion from training data
- Barriers in authentication or identity verification

2. Bias, Discrimination, and Inequitable Outcomes

Risks where AI systems produce unfair or harmful outcomes.

- Discriminatory outputs
- Use of proxy variables that encode disability bias
- Unequal error rates across groups
- Reinforcement of structural inequities

3. Safety, Reliability, and Performance Failures

Risks related to unpredictable or unsafe system behavior.

- Unreliable predictions
- Failures in safety-critical contexts
- Model drift affecting accessibility
- Unstable assistive or adaptive outputs

4. Data Governance and Privacy Risks

Risks involving misuse, exposure, or improper handling of data.

- Exposure of sensitive disability-related data
- Lack of informed consent or transparency
- Insecure data flows
- Misaligned data retention practices

5. Human Oversight and Accountability Gaps

Risks arising from weak governance or insufficient human control.

- Absence of human-in-the-loop safeguards
- Unclear accountability roles
- Inadequate escalation mechanisms
- Insufficient monitoring and review

6. Systemic and Societal Harms

Risks that reinforce exclusion or limit participation in society.

- Reduced access to public services
- Barriers to employment or education
- Loss of autonomy or dignity
- Reinforcement of stigma

7. Interoperability and Data Privacy Impact-Related Risks

Risks that arise when AI systems interact with digital public infrastructure.

- Failure in identity verification
- Exclusion from payments or benefits
- Misalignment with health or mobility systems
- Accessibility breakdowns across systems

8. Emerging and High-Risk AI Behaviors

Risks linked to advanced or unpredictable AI capabilities.

- Hallucinations that affect accessibility
- Manipulative or coercive outputs
- Autonomous decision-making without safeguards
- Unpredictable multimodal interactions

Risk Codes and Traceability

Each risk is assigned a unique code (for example, EA3 or BD2). These codes are used consistently across:

- Audit framework checklists
- Benchmark indicators
- Safety and red-teaming protocols
- Scenario testing
- Validation outputs
- Governance maturity assessments

This coding system ensures that risks identified in one stage can be tracked throughout the entire governance lifecycle.

When to Use the Taxonomy

The taxonomy can be applied at multiple stages:

- During audits: Identify relevant risk categories
- During scenario testing: Map observed behaviors to risk codes
- During governance reviews: Track mitigation actions by category
- During DPI integration: Identify cross-system risks (IP category)
- During expert review: Ensure alignment with global standards

Harms Repository

A Harms Repository is intended to complement the Risk Taxonomy. The Harms Repository is a centralized collection of documented harms, risks, incidents, and negative impacts associated with a product, process, technology, or organization. It is often used in AI governance, risk management, cybersecurity, healthcare, and product development to help teams identify, track, and mitigate potential harms.

The repository should be a structured, continuously updated database that documents AI-related harms affecting persons with disabilities and older persons. It provides a unified and comparable foundation for risk assessment, regulatory compliance review, and system-level accountability.

A governance-aligned Harm Repository should function as a living system with structured opportunities to capture newly emerging or previously missed harms through continuous intake, periodic review cycles, and feedback loops tied to monitoring, incident reporting, and cross-system comparisons. This approach maintains completeness, supports accountability, and ensures the repository evolves with real-world use, and supports the identification of new and previously undocumented harms through contributions from persons with disabilities, researchers, civil society organizations, and public agencies. Emerging harms should be reviewed periodically and incorporated into future versions of the taxonomy where appropriate.

Purpose

- Capture real-world harms that technical testing alone cannot detect.
- Provide developers, policymakers, DPI teams, and regulators with actionable evidence.
- Ensure that disability-inclusive harms are visible, traceable, and systematically addressed.

What the Repository Should Include

- Documented harms from AI systems across public services sectors (health, employment, education, justice, social protection, transportation, etc.).
- Evidence from lived experience, civil society reports, academic research, and regulatory findings.
- Metadata for each harm, including population affected, system type, severity, frequency, and context.
- Links to relevant standards, safeguards, and mitigation strategies.

Appendix 2: Use Case Scenarios

Future versions of this framework may incorporate quantitative measures, including error rates, complaint volumes, accessibility testing results, and documented incidents, to complement the illustrative scenarios presented in this appendix.

2.1 Telehealth AI Misinterprets AAC Speech

A national telehealth platform uses an AI-driven symptom triage chatbot. The model misinterprets AAC-generated speech, resulting in inaccurate symptom summaries and unsuitable care recommendations. Some AAC users are incorrectly flagged as incoherent, which delays their access to appropriate care.

Risk taxonomy

- Misclassification risk: Incorrect interpretation of AAC speech.
- Safety-critical risk: Potential for incorrect triage severity.
- Accessibility failure risk: System not designed for AAC inputs.
- Exclusion risk: AAC users unable to complete triage.
- Dignity and autonomy risk: Users labeled as incoherent.
- Systemic bias risk: AAC users are systematically disadvantaged.

Audit framework

- Data audit: No AAC speech represented in training data.
- Model audit: No robustness testing for AT-generated speech.
- Interface audit: Chatbot lacks AAC-compatible input pathways.
- Accessibility audit: No alternative input modes.
- Oversight audit: No escalation for uncertain outputs.
- Monitoring audit: No AAC-specific error tracking.

Benchmark application

- AAC speech recognition benchmark: High error rate.
- Cognitive load benchmark: Instructions too complex.
- AT interoperability benchmark: Fails to accept AAC inputs.

Safety protocol

- Misclassification safety test: Incorrect triage severity.
- Urgency safety test: Failure to detect uncertainty.
- AT failure mode test: No fallback mechanism.

Harms repository

- Health access harm: Delayed or incorrect care.
- Diagnostic delay harm: Misinterpretation leads to delay.
- AT-related harm: AAC incompatibility.
- Dignity harm: Users labeled as incoherent.

Governance maturity

There is low maturity in inclusive data practices, AT interoperability, safety assurance, accessible escalation pathways, and monitoring and feedback mechanisms.

DPI-AI interoperability

There are failures in the DPI accessibility layer, AT interoperability layer, and safety and escalation layer.

2.2: Voice Authentication Rejects Older Adults

A national digital ID system uses voice authentication for account access. Older adults face repeated authentication failures due to age-related vocal changes, including pitch variation, tremor, and breathiness. After several failed attempts, the system locks users

out, blocking access to essential services such as health care, banking, and social protection.

Risk taxonomy

- Misclassification risk: Incorrect rejection of legitimate users.
- Accessibility failure risk: System not designed for age-related vocal variation.
- Exclusion risk: Older adults are disproportionately affected.
- Systemic bias risk: Training data and model behavior reinforce age bias.
- Identity and access risk: Authentication failures compromise access to identity-linked services.
- Service denial risk: Lockouts block access to critical public and private services.

Audit framework

- Data audit: Training data lacks adequate representation of older adults' voices.
- Model audit: No robustness testing for age-related vocal changes.
- Interface audit: No alternative authentication pathways provided.
- Accessibility audit: Absence of multimodal fallback options.
- Oversight audit: No human review or escalation for repeated failures.
- Monitoring audit: No tracking of age-related error patterns or disparities.

Benchmark application

- Age-related voice variation benchmark: High false-reject rate for older adults.
- Biometric fairness benchmark: Significant performance disparities across age groups.
- Multimodal access benchmark: System fails to provide alternative authentication modalities.

Safety protocol

- Identity safety test: Incorrect rejection of legitimate users.
- Fallback safety test: No safe alternative authentication method.
- Lockout safety test: System escalates to account freeze without human intervention.

Harms repository

- Access denial harm: Users are unable to access essential services.
- Identity harm: Authentication failures undermine identity verification.
- Financial access harm: Banking and payment systems become inaccessible.
- Social protection access harm: Benefits and entitlements are blocked.
- Dignity harm: Repeated failures undermine user confidence and autonomy.

Governance maturity

Inclusive biometric data practices, multimodal authentication readiness, identity-system safety assurance, monitoring of demographic disparities, and escalation and human-in-the-loop pathways are immature.

DPI-AI interoperability

This scenario impacts the identity, payments, health access, and social protection layers.

2.3: Navigation App Routes Wheelchair Users Through Inaccessible Paths

A widely used navigation app provides pedestrian routing guidance. Wheelchair users report that the system repeatedly directs them through staircases, steep slopes, narrow passages, and other inaccessible paths. The model lacks accessibility metadata and offers no fallback routes. As a result, users face delays, safety risks, and, in some cases, physical harm.

Risk taxonomy

- Misclassification risk: Incorrectly labeling inaccessible paths as accessible.
- Safety-critical risk: Directing users toward physical hazards.
- Accessibility failure risk: No wheelchair-accessible routing logic.
- Exclusion risk: System not usable by wheelchair users.
- Systemic bias risk: Accessibility needs are not represented in model design.
- Service denial risk: Users cannot rely on the system for safe navigation.

Audit framework

- Data audit: No accessibility metadata, such as curb cuts, ramps, slopes, or obstacles, in training data.
- Model audit: No testing for wheelchair-accessible routing.
- Interface audit: No option to specify mobility needs.
- Accessibility audit: No alternative route suggestions.
- Oversight audit: No human review of accessibility complaints.
- Monitoring audit: No tracking of accessibility-related routing errors.

Benchmark application

- Mobility hazard detection benchmark: Fails to detect stairs, steep slopes, and narrow passages.
- Environmental accessibility benchmark: No integration of accessibility layers.
- AT interoperability benchmark: No compatibility with wheelchair navigation tools.

Safety protocol

- Hazard safety test: Routes users toward unsafe environments.
- Fallback safety test: No safe alternative route provided.
- Uncertainty safety test: System does not flag low-confidence routing.

Harms repository

- Physical safety harm: Exposure to unsafe terrain.
- AT-related harm: Wheelchair users cannot navigate safely.
- Mobility access harm: Barriers to independent travel.
- Dignity harm: Systemic exclusion from basic mobility.

Governance maturity

Accessibility data governance, integration of accessibility metadata into national mapping systems, safety assurance for mobility-critical applications, monitoring and escalation pathways for accessibility complaints, and cross-agency coordination among transport, urban planning, and digital services all exhibit low levels of maturity.

DPI-AI interoperability

This scenario reveals failures in the mobility, urban infrastructure, and safety and escalation layers.

Bibliography

Disability-Specific Taxonomies of Risks

1. **Wang, Kameswaran & Kacorri (2025).** *Toward a Taxonomy of Algorithmic Harms for Disability: A Systematic Review*. AIES 2025, 8(3), 2649–2665.
<https://doi.org/10.1609/aies.v8i3.36745> ([doi.org in Bing](#)). Systematic review of 76 papers across HCI, accessibility, and responsible AI venues, annotating 175 harm instances using the Shelby et al. framework and identifying disability-specific harm patterns across individual, community, and societal levels.
2. **DisaBench (2026).** *A Benchmark for Disability-Related Harms in Language Models*. arXiv:2605.12702. Co-created through red-teaming and workshops with over 100 people with disabilities, organizing twelve harm categories into five top-level groups with three severity dimensions. Identifies failure modes such as pity framing, harmful compliance, reinforcement of negative behaviors, and quality degradation when disability is disclosed. Preprint; not yet peer-reviewed.

General Frameworks Underpinning Disability-Focused Harm Taxonomies

3. **Shelby et al. (2023).** *Sociotechnical Harms of Algorithmic Systems: Scoping a Taxonomy for Harm Reduction*. AIES '23, 723–741. Foundational five-theme harm taxonomy (representational, allocative, quality-of-service, interpersonal, societal) used as the structural backbone for disability-specific harm classification.
4. **Hoffmann & Frase (2023).** *CSET AI Harm Framework*. Center for Security and Emerging Technology. Provides incident-oriented harm categories including physical safety, financial loss, property damage, detrimental content, bias, and rights violations, distinguishing realized from potential harms.
5. **AIAAIC (2024).** *Human-Centred Harms Taxonomy*. arXiv:2407.01294. Broad incident-driven taxonomy with nine top-level and sixty-nine sub-categories derived from over 1,500 documented AI incidents.

Governance-Aligned Frameworks Relevant to Traceability, Lifecycle Management, and Harm Repositories

6. **OECD (2024).** *Framework for the Classification of AI Incidents*. Organisation for Economic Co-operation and Development. International schema for categorizing AI incidents by harm type, severity, and contextual factors; supports alignment of harm repositories with global reporting norms.

7. **UNESCO (2021).** *Recommendation on the Ethics of Artificial Intelligence*. United Nations Educational, Scientific and Cultural Organization. Establishes human-rights-based principles including accessibility, non-discrimination, and societal harm prevention, providing normative grounding for disability-inclusive AI governance.
8. **NIST (2023).** *AI Risk Management Framework (AI RMF 1.0)*. National Institute of Standards and Technology. Lifecycle-based risk management framework emphasizing identification, measurement, mitigation, and continuous monitoring; directly supports mechanisms for updating and retiring risk codes.
9. **G7 (2023).** *Hiroshima Process International Guiding Principles for AI*. Group of Seven. Cross-national principles emphasizing safety, accountability, and harm mitigation, situating disability-specific harms within broader global governance expectations.
10. **Partnership on AI (ongoing).** *AI Incident Database*. Large, curated repository of real-world AI harms and failures; complements AIAAIC and provides additional grounding for incident-driven harm identification.
11. **UN Committee on the Rights of Persons with Disabilities (2014).** *General Comment No. 2 on Accessibility*. Foundational human-rights guidance on accessibility obligations under the CRPD; supports the normative basis for disability-aligned harm taxonomies and traceability requirements.
12. **European Union (2024).** *AI Act — Final Text*. Regulatory framework including provisions on accessibility, high-risk systems, incident reporting, and post-market monitoring; provides statutory grounding for traceability and harm-repository design.