

**INTERSESSIONAL PANEL OF THE UNITED NATIONS COMMISSION
ON SCIENCE AND TECHNOLOGY FOR DEVELOPMENT (CSTD)**

**Geneva, Switzerland
6 November 2026**

**Contribution by ODET
to the CSTD 2026-2027 priority theme on
“Data Governance for Development: From Asymmetry to Equitable Stewardship”**

DISCLAIMER: The views presented here are the contributors' and do not necessarily reflect the views and position of the United Nations or the UN Trade and Development.



Submission to the CSTD 30 Request for inputs

Data Governance for Development: From Asymmetry to Equitable Stewardship

The United Nations Office for Digital and Emerging Technologies (ODET) welcomes the opportunity to contribute to the CSTD issue paper themed '*Data Governance for Development: From Asymmetry to Equitable Stewardship*'. This submission responds to the seven questions of the request for inputs, drawing principally on two multistakeholder convenings organized by ODET with partners in 2026:

- From Practice to Policy: Data Commons for Equitable Data Governance Frameworks (20 February 2026), convened in direct support of the [CSTD Multi-stakeholder Working Group on Data Governance at All Levels](#) (established pursuant to General Assembly [resolution 79/1](#)). Following the Working Group's initial meetings, Member States requested further practical insight from ongoing initiatives worldwide; the workshop responded by convening Working Group members and observers alongside operators of live Data Commons initiatives.
- Data Governance for the Public Good: Equity, Openness and Community Stewardship (Open Source Week, 2026), which examined the conditions under which affected communities participate in data governance as co-creators and stewards rather than as data subjects.

It is further informed by ODET's previous work on a Data Commons implementation framework, developed with regional consultations in Bangalore, India (January 2025), and Pretoria, South Africa (April 2025).

This submission also builds on the wider trajectory of United Nations digital cooperation policy that has progressively placed data governance at its centre: the 2019 report of the Secretary-General's High-Level Panel on Digital Cooperation, [The Age of Digital Interdependence](#); the Secretary-General's 2020 [Roadmap for Digital Cooperation](#); the 2021 [Our Common Agenda](#) report; the 2024 final report of the High-Level Advisory Body on Artificial Intelligence, [Governing AI for Humanity](#); the [Global Digital Compact](#), adopted at the 2024 Summit of the Future, which tasked the CSTD with establishing the Working Group on Data Governance at All Levels, and the [Preliminary Report](#) of the Independent International Scientific Panel on AI.

ODET is pleased to contribute to this progression of dialogue through the present submission.

Question 1. Key challenges in building and leveraging data capacities

Interoperability barriers are as significant within countries as across borders: Interoperability remains one of the most significant barriers to unlocking the developmental value of data. While international discussions often focus on cross-border data sharing, participants in both workshops emphasized that equally consequential barriers exist within countries themselves. Government ministries, public agencies and sectoral institutions frequently operate through disconnected systems, incompatible standards and fragmented administrative practices that limit the ability to share, combine, and reuse data effectively.

As a result, data that could support better public services, research, innovation and evidence-based policymaking often remain siloed across institutions.



Submission to the CSTD 30 Request for inputs

Data Governance for Development: From Asymmetry to Equitable Stewardship

The experience of the [Amazon Rainforest Evolution Index](#), led by the [CERTI Amazônia Institute](#), illustrates the scale of this challenge. The initiative has spent years integrating data on deforestation, land use, river systems, and forest traceability from multiple sources and institutions. Participants noted that a substantial portion of this effort has involved harmonizing heterogeneous datasets, adapting to changing formats, and maintaining technical mechanisms that allow information to remain interoperable over time. The lesson is that interoperability is not merely a technical issue. It is a continuous governance and coordination challenge requiring alignment of standards, incentives, institutional practices, and long-term investment. Without such alignment, the potential developmental benefits of data remain difficult to realize even where political commitment and data availability already exist.

Foundational data does not exist for many populations, and sovereignty presupposes infrastructure : The [Malawi Voice Data Commons](#) had to construct training corpora from near zero, collecting and standardizing thousands of voice recordings in low-bandwidth environments before the digital emergency service notifications could be built. This example demonstrates how commitments to control and govern local data require corresponding in-country capacities, from storage to processing. Zooming out further, at present Africa accounts for well under one per cent of global data centre capacity; [ITU records](#) 2.2 billion people offline in 2025, 96 per cent of them in low- and middle-income countries, with rural internet use at 58 per cent against 85 per cent in urban areas. If governance commitments outrun infrastructure realities, their effect remains on paper.

Public data is often inaccessible in practice despite legal entitlement: Participants in the Pretoria consultation reported that data scientists cannot obtain basic public datasets such as urban plans; that government documents are published almost exclusively as PDFs, with requests for machine-processable formats routinely refused despite national mandates; and that access decisions turn on who is asking rather than why and how data will be used.

Benefit-sharing is constrained by an unresolved traceability problem: Benefit-sharing remains difficult to operationalize in practice. Participants emphasized that much of the value derived from data emerges when datasets are combined, reused, and applied across multiple contexts, including in AI systems and commercial digital services. Yet it remains difficult to determine how individual datasets or data contributors have contributed to downstream outcomes, particularly after data has been aggregated, transformed or incorporated into larger systems. This attribution challenge complicates efforts to design fair and transparent benefit-sharing arrangements and remains a significant obstacle to translating broad commitments into operational mechanisms.

Data commons offer one potential response to this challenge: By establishing agreed governance frameworks, stewardship arrangements, transparency requirements and clear rules



Submission to the CSTD 30 Request for inputs

Data Governance for Development: From Asymmetry to Equitable Stewardship

on access and reuse, data commons can help create more equitable approaches to participation and value-sharing. While they do not fully resolve the attribution challenge, they can provide institutional mechanisms through which the benefits derived from data are governed and shared in ways that better reflect collective contributions and public-interest objectives.

Sustainability beyond project funding is the recurring limitation for data commons: Across all cases examined, the transition from grant-funded pilot to durable national infrastructure emerged as the decisive risk, requiring institutional embedding and long-term financing rather than renewed project cycles.

Question 2. How inequalities affect the generation, collection, governance and use of data

Exclusion is frequently designed into systems before any data is collected: The Malawi case is illustrative. Conventional emergency reporting systems assume literacy, smartphone access and command of a dominant language; yet a large share of Malawi's rural population meets none of these assumptions. The consequence is not merely that these populations are underserved; they do not even appear in the data that subsequently informs resource allocation. Inequality in data generation cascades into inequality in policymaking and service delivery. The absence of certain populations is not visible to the systems that produce it.

Linguistic underrepresentation compounds this at scale: Africa accounts for under 0.5 per cent of contributions to machine learning and large language model development. Participants in Pretoria cited severe mistranslations as a direct cost of poor language representation, noting that language inclusion is comparatively inexpensive yet routinely omitted. As underrepresented languages remain absent from training data, systems built on that data entrench the exclusion.

The gender data gap follows the connectivity gap: [ITU records](#) 77 per cent of men using the internet against 71 per cent of women in 2025, some 280 million more men than women online. Populations underrepresented in digital life are correspondingly underrepresented in the administrative and behavioral data generated through digital services, with consequences for the evidence base of policy in developing and developed countries alike.

Individual consent frameworks fit poorly where knowledge is held collectively : Participants in Pretoria emphasized that many African communities hold knowledge communally rather than individually, making communal consent essential to legitimate sharing. Instruments built exclusively on individual consent either exclude such knowledge or authorize its extraction without the agreement of those who hold it.

Measures under way: Deliberate curation for representativeness is emerging as a response. The [Institutional Data Initiative](#) at Harvard Law School Library curates datasets to improve balance and diversity in AI training data; the [LINGO Open Call](#) supports low-resource European



Submission to the CSTD 30 Request for inputs

Data Governance for Development: From Asymmetry to Equitable Stewardship

languages, with a recently announced expansion to African languages funding dataset creation and technical capacity; and [Mozilla Common Voice](#) demonstrates a community-governed model in which contributors also validate and use the resulting datasets. Representative datasets are achievable, resource-intensive, and dependent on sustained rather than one-off financing.

Question 3. National strategies, regulatory frameworks and policy instruments: lessons learned

Enactment without administrative capacity produces rights on paper: South Africa's [Promotion of Access to Information Act \(2000\)](#) forbids officials from requiring requesters to justify their need for access. Participants reported that this is regularly violated in practice and that the framework is both poorly implemented and widely misunderstood within government itself. Access rights require procedure, official training and enforcement to become operative.

Formal mandates without default standards and compliance mechanisms do not change practice: Notwithstanding national mandates, government reporting in South Africa continues to lack standardized machine-readable formats. Machine-processable publications would enable innovation in local languages.

Definitional carve-outs adopted before AI now generate governance gaps. India's [Digital Personal Data Protection Act \(2023\)](#) excludes publicly available personal data from its scope. Participants in Bangalore raised the unresolved question of whether such data properly belongs in shared data environments and what protections are required in relation to AI training. Exclusions that were reasonable when the Act was drafted require revisiting, given current practice.

AI policy developed apart from constitutional and data protection traditions risks incoherence: Participants noted that South Africa's draft AI policy makes minimal reference to constitutional values such as dignity and privacy, and that fragmented national and regional approaches hinder coordinated governance.

Positive direction: Two approaches attracted broad support. The first is graded rather than binary openness: the appropriate position depends on context, rights, risks and incentives, and while some data should be open by default, community-held knowledge should remain under community control, with those boundaries transparent, fair and enforceable. The second is purpose limitation adapted to shared data environments, which participants in Bangalore suggested could be advanced through a United Nations-led mechanism.

Question 4. Initiatives reducing asymmetries and promoting equitable data stewardship

Public services and emergency response, Malawi Voice Data Commons: A toll-free, voice-based reporting system accessible from basic mobile phones, with reports transcribed by locally trained speech recognition models and routed to district-level responders. Its most



Submission to the CSTD 30 Request for inputs

Data Governance for Development: From Asymmetry to Equitable Stewardship

transferable governance feature is the separation of data classes: long-term voice training datasets are handled distinctly from short-term emergency reporting data, which is typically deleted after use to protect privacy, a replicable privacy-by-design norm for service delivery in low-connectivity settings. The remaining challenge is institutional embedding to carry the system from grant-funded projects to national infrastructure.

Environmental governance, Amazon Rainforest Evolution Index (CERTI Amazônia Institute): Two decades of work integrating and standardizing fragmented environmental data into interoperable systems, now supporting AI-ready tools and a domain-specific large language model focused on Amazon data that generates trackable indices over time. The case demonstrates the value of contextualized AI built on curated, region-specific datasets over generic models lacking local accuracy. Data quality, reliability, and continuous updating require sustained monitoring, processing capacity, and infrastructure investment.

Licensing that retains value within local ecosystems: Participants in Pretoria described licensing models giving communities control over how their data is used. A South African automatic speech recognition dataset is freely available to African researchers while other users pay a fee reinvested into local data creation. Related instruments discussed included the [Nudu Open License](#), developed with community input to guide ethical open data sharing. These approaches address benefit-sharing at the point of access rather than attempting downstream attribution, and merit closer examination by the Commission for that reason.

Private-sector and multistakeholder contributions: Participants highlighted the important role that private actors can play in expanding access to data, infrastructure, and technical expertise that support development outcomes. Examples include [Microsoft's Open Data Campaign](#), which supports initiatives such as the [Institutional Data Initiative \(IDI\)](#) and [CORE](#) to improve access to high-quality and diverse datasets and research infrastructure for AI innovation. Similar efforts can be seen in multistakeholder initiatives such as the [Digital Public Goods Alliance](#) and the [FAO Hand-in-Hand Geospatial Platform](#), which facilitate wider access to data and digital resources for sustainable development. Participants also pointed to the growing influence of community-centred governance frameworks, including the [CARE Principles for Indigenous Data Governance](#), which emphasize collective benefit, authority to control, responsibility and ethics in the stewardship of data. Together, these examples demonstrate that reducing data asymmetries requires not only investment and technical capacity, but also governance approaches that ensure affected communities retain agency over how data is accessed, used and shared. These experiences suggest that external support can strengthen local innovation ecosystems and public-interest outcomes, provided that stewardship responsibilities, access conditions and benefit-sharing arrangements are clearly defined, particularly where data contributes to commercially valuable AI systems.



Submission to the CSTD 30 Request for inputs

Data Governance for Development: From Asymmetry to Equitable Stewardship

Question 5. Mechanisms and institutional arrangements fostering collaboration

The data fiduciary model: Participants in both Bangalore and Pretoria proposed a data fiduciary as a neutral intermediary between providers and users, capable of moderating extraction and ensuring fair return, citing [AI4Bharat](#) which generated significant benefit from public language data without reciprocal value exchange as illustrating the gap. Workshop participants separately pointed to academic peer review as an existing governance model for reliability and accountability that could be adapted to dataset curation, noting that poorly-curated data produces errors and hallucinations in AI systems.

Convening practitioners alongside policymakers: The February 2026 workshop itself constitutes a replicable mechanism: it was convened because Member States, following the Working Group on Data Governance's initial meetings, requested practical insight from ongoing initiatives, an indication that principles-level deliberation had outrun the available operational evidence.

Question 6. Bilateral, regional and international partnerships

Regional environmental data cooperation: The Amazon experience illustrates two complementary dimensions. Horizontal cooperation pools environmental data across multiple Amazonian countries, requiring alignment of standards, legal frameworks and policies to permit cross-border exchange. Vertical integration operates within a single jurisdiction across levels of government, as with the integration of data across local, state and national levels within Brazil's Legal Amazon. The benefit is a coherent evidence base for regional environmental governance; the cost lies in sustaining standard alignment across jurisdictions with differing legal traditions and administrative capacities.

Public-private partnerships as enablers of data ecosystems: Participants highlighted that partnerships between public institutions, private companies, academia, and civil society can help address persistent gaps in infrastructure, financing, and technical capacity. Such arrangements can accelerate the development of datasets, digital infrastructure, and AI capabilities that would otherwise remain beyond the reach of many institutions in developing countries. However, their effectiveness depends on clear governance arrangements established from the outset, including rules on data access, reuse, stewardship, and benefit-sharing. Several participants stressed that public trust and long-term sustainability are strengthened when partnerships are designed to build local capacity and institutional ownership rather than creating continued dependence on external providers.

Question 7. Strengthening international cooperation, and the contribution of the CSTD

ODET offers the following recommendations; each grounded in the feedback and research carried out to support the convenings of 2026:



Submission to the CSTD 30 Request for inputs

Data Governance for Development: From Asymmetry to Equitable Stewardship

Support trusted stewardship mechanisms for data sharing: Effective data governance requires institutions capable of building trust between data providers, users and affected communities. The CSTD may wish to encourage further exploration of data fiduciaries, trusted intermediaries and similar stewardship arrangements that can facilitate data sharing, oversee compliance with agreed rules, and help ensure that benefits are distributed fairly across participants.

Create stronger incentives for participation in data ecosystems: Many stakeholders remain reluctant to share data where the benefits of participation are unclear or unevenly distributed. Governance frameworks should therefore incorporate mechanisms that recognize and reward contributions, including through access to shared resources, capacity-building, technical support, improved public services, or other forms of value creation. Sustainable participation depends on stakeholders perceiving clear and equitable returns from data sharing.

Finance foundational datasets and data infrastructure as public goods: Building representative datasets for underrepresented languages, communities and development contexts requires sustained investment and long-term maintenance. International cooperation should support the transition from short-term project funding to durable institutional arrangements that treat foundational datasets, standards and supporting infrastructure as strategic public assets.

Recognize communal consent and community stewardship in data governance frameworks: Approaches based exclusively on individual consent are often insufficient where knowledge, cultural heritage, or data are held collectively. International discussions should continue exploring governance mechanisms that recognize community-held knowledge, collective interests and stewardship responsibilities, while ensuring affected communities retain meaningful agency over how their data is accessed, used and shared.

Promote practical data governance models through continued multistakeholder engagement: The experience of the Working Group on Data Governance and the February 2026 workshop demonstrated the value of bringing practitioners, policymakers, researchers, and communities into the same conversation. Establishing recurring channels for operational evidence and practical experience would help ensure that future recommendations remain grounded in implementable governance models and real-world development needs.

One conclusion recurs across this evidence: advancing equitable data governance requires moving beyond abstract principles toward implementable institutional models. The CSTD's priority theme is well framed to make that move, and ODET stands ready to support it, including through continued multistakeholder convening.



Submission to the CSTD 30 Request for inputs

Data Governance for Development: From Asymmetry to Equitable Stewardship

Contact persons

Isabel de Sola Criado — isabel.desolacriado@un.org

Francesco Stabilito — francesco.stabilito@un.org

Thierno Ousmane BA — thierno.ba@un.org

Sources

- *Report of the ODET Workshop in Support of the Working Group on Data Governance, "From Practice to Policy: Data Commons for Equitable Data Governance Frameworks", 20 February 2026*
- *Report Summary, Open-Source Week 2026 Workshop, "Data Governance for the Public Good: Equity, Openness and Community Stewardship", 2026*
- *ODET, Data Commons White Paper: Implementing equitable and trustworthy data governance (internal draft).*