G20 Data Governance Priorities for Equitable and Sustainable Artificial Intelligence

South Africa's G20 Presidency has recognised the critical relationship between data governance and Artificial Intelligence (AI) in the establishment of the High-Level Task Force on AI, Data Governance and Innovation for Sustainable Development. Its call for G20 commitment to global collaboration reflects the Presidency's understanding that the way data is governed will significantly determine whether AI will amplify existing inequalities, as current evidence on the development of advanced data-driven technologies indicates, or whether it can be harnessed for the benefit of all humanity. This requires a nuanced understanding of the interplay between data governance and the rapidly evolving forms of Artificial Intelligence (AI) as the latest general-purpose technology, from Large Language Models (LLMs) to Generative AI and Agentic AI.

This was the purpose of the South African G20 Presidency's Data Dialogue, which provided an opportunity for G20 members to engage with United Nations, multilateral and regional organisations and T20 on the dynamic evolution of data governance globally, regionally and nationally, increasingly in the context of seeking to create rights-preserving, safe, secure and trustworthy AI ecosystems and to forge pathways towards a more just and sustainable digital future for all

Data governance is a recurring priority across several global and regional frameworks, including the Global Digital Compact (GDC), the World Summit on the Information Society (WSIS) +20 Review, various UN multilateral organisation responses as well as regional initiatives, in the African Union, Association of Southeast Asian Nations, and European Union, and other multilateral organisations such as OECD, the World Bank and the African Development Bank. This annex draws on this dynamic repository of knowledge and experience on data governance for AI, together with the rich discussions in the high-level workshop and contributions to the Data Dialogue Issues Note.

Building on the Brazilian G20 surfacing of data governance a critical issues for contemporary society and economy this annex serves as a guide to assist G20 members in framing their approaches to the governance of data as the central input of rapidly evolving AI technologies; to limit the amplification of inequalities and exclusions, perpetuation of collective and systemic harms associated with large-scale data driven technologies concentrated in the hands of few actors; and to complement on-going efforts to create a trustworthy system of global, regional and national data governance frameworks to enable AI to serve all of humanity.

Seven key issues arose from the discussions on data governance for equitable and sustainable AI:

1. Data Value as Driver of AI: The circulation of data is a critical aspect of data value creation and a driver of the digital economy. Previous G20 Presidencies committed to innovative, interoperable, and inclusive mechanisms enabling data flow with trust, respecting data protection and privacy. Data gains value through processing, sharing, and reuse across contexts. In AI, data quality, representativeness, and scope directly shape model performance, fairness, robustness, and compliance. A clear, context-specific approach to data valuation, adapting existing models to AI's dynamic reuse and recombination of data and creating an enabling environment of public and local private value creation, is essential. This can enhance accountability and support proportional regulation, ensuring fair value creation and distribution.

2. Integrated and Interoperable Data Governance Approaches: How data is governed is a key determinant of AI's promise for innovation and sustainable development, especially in developing economies. An integrated data governance framework for AI should include data standards, interoperability measures, accessibility protocols, quality assurance, data-sharing agreements (including cross-border flows), and robust security and privacy guidelines. Innovative, standardised, and integrated interoperable

data systems that enable trusted data flow within and across borders are essential for high-quality, globally representative datasets that reflect diversity in language, culture, race, and geography.

3. Data Justice: Data Justice requires moving beyond the narrow focus on individual rights in dominant data governance frameworks, to recognise collective rights and systemic inequalities, and calls for fairer access and use of data, computing power and opportunities for digital literacy. Data justice advocates for addressing persistent and growing digital inequalities, including by upholding economic rights to data, ensuring inclusive and democratic participation in data governance, promoting equitable access to data and digital resources, preventing exploitative and anti-competitive practices, and enabling meaningful redress where harms occur. For Instance, AUDPF outlined that economic rights to data must be recognised alongside efforts to ensure equitable opportunities in the broader data economy.

4. Data Commons: Data commons establish a systemic framework that acknowledges data's dynamic nature and multiplied value through sharing and agile stewardship. They offer a structured approach to creating equitable, well-governed data ecosystems where diverse stakeholders contribute, access, and utilise data. In economic terms, data commons help correct imbalances created by decades of supply-side, market-led resource allocation, driven by profit maximization. Common-based approaches, such as open and shared data repositories, data lakes, and mandated access to data in the public interest, enable broader access and public benefit by meeting demand-side constraints, such as the affordability of critical inputs, like data. Such approaches include transparent and accountable governance structures, protocols, and interoperability standards into an AI ecosystem that recognises that data's true potential arises from dynamic interactions between stakeholders, processes, and technological infrastructures.

5. Emerging challenges for data governance evolving AI systems: Developments in AI introduce novel challenges for data governance. Agentic AI, which acts independently across various datasets and cloud platforms, raises novel concerns about privacy, accountability and data traceability. These systems make it harder to apply key data protection principles like data minimisation and purpose limitation. At the same time, the rapid growth of commercial data markets has raised questions about data ownership, fair compensation and legal rights. Generative AI is also making it easier to work with unstructured data, which has blurred the lines between personal and non-personal information, particularly as de-anonymisation becomes more feasible. Limited

human oversight increases the risk of opaque decision-making—so-called "black box" systems—that are difficult to regulate and threaten to worsen existing social and economic inequalities. Another emerging issue is the growing use of synthetic data, especially in training models developed by generative AI. While useful, synthetic data cannot fully replace real-world data. Over-reliance may reduce model accuracy and lead to errors or false outputs, often referred to as "hallucinations".

6. Environmental Impact of Data-Driven Technologies: The growing demand for data transmission, processing, and storage has increased energy consumption and carbon emissions. The ICT sector emitted an estimated 0.69 to 1.6 gigatons of CO_2 equivalents, accounting for 1.5% to 3.2% of global GHG emissions in 2020. Additionally, the production and use of digital devices, data centres, and ICT networks contribute 6% to 12% of global electricity use. Mitigating these impacts requires policy reforms, technological innovations, and stakeholder action to make business models more circular, energy-efficient, and responsible for consumption. UNCTAD advocates for a global shift towards a circular digital economy, prioritising recycling, re-use, and recovery of digital materials, resource optimisation, stricter environmental standards, investment in renewable energy, and international cooperation.

7. Regional Harmonisation: Several regional organisations such the African Union and the European Union have initiatives and frameworks that aim at enhancing the harmonisation of data governance frameworks and enabling the deployment of AI systems locally while mitigating risks associated with global AI companies' operating across jurisdictions.

8. Multilateral and Multi-stakeholder Cooperation: The multidimensional and transnational nature of data necessitates multilateral and multi-stakeholder cooperation for developing functional global governance frameworks to maximise and equitably distribute AI gains and address environmental impacts. Realising global digital public goods, such as data and their governance at the national and regional level, requires global responsibility for the resourcing and governance of them. The G20, in line with its commitments to rights-respecting and just AI, must actively participate in, shape, and guide these efforts—helping to build inclusive, accountable, and fair global data and AI governance systems.

The G20 data dialogue underscored the importance of data governance for equitable, inclusive, and just AI outcomes, along with practical steps for policymakers. It highlighted the persistent, evolving and emerging challenges in this space and explored

how to address these, including by leveraging existing governance frameworks. Priority actions include:

- Support the development of regional data commons, with investment in inclusive, culturally and linguistically relevant datasets and training material, particularly across Majority World nations.

- Require mandatory data audits for public AI systems and promote binding standards for algorithmic accountability and explainability, aligned with international human rights law.

- Explore progressive data taxation and benefit-sharing schemes to ensure companies extracting value from national datasets contribute to local digital development and AI capacity-building.

- Embed environmental consideration into data governance policies to address the ecological impact of data-driven AI technologies.

- Promote harmonised standards, cross-border data flow agreements, and South-South cooperation to reduce global data asymmetries and encourage inclusive participation in governance fora.

- Recognise the common and different uses of data across AI systems and ensure that governance strategies are tailored accordingly.

- Recognize the importance of keeping data governance abreast of AI advancements and the need to adapt a gradual, interoperable, and adaptive approach to global data governance that promotes trust and responsible data exchange while respecting national contexts.

- Promote transparent and accountable data management throughout the entire AI value and encourage data lifecycle-based governance frameworks.

While structural inequalities in the global digital economy remain, targeted reforms and coordinated international action can drive meaningful progress. G20 member states reaffirm their commitment to strengthening data governance in the public interest and to upholding all human rights, including those linked to economic justice and environmental sustainability. Going forward, governance should ensure that all countries and communities can exercise control, sovereignty, and ownership over their data across the entire data and AI value chain.

Solidarity   Equality   Sustainability

The South African Presidency thanks the AUC, UN ODET, and RIA for facilitating the Data Dialogue and UNCTAD, UNESCO, OECD, Data Privacy, Open Up, African Development Bank, ILO and the United Arab Emirates for their valuable input into the workshop and the Issue Paper presented at the third meeting of the AI Task Force, which forms the basis of this annex.