Working Group Report on Data Governance in the Digital Age

Data Governance Toolkit

Navigating Data in the Digital Age

July 2025

BROADBAND COMMISSION
FOR SUSTAINABLE DEVELOPMENT

ITU

unesco

CI/DIT/2025/DT/01

Data Governance Toolkit

Navigating Data in the Digital Age

July 2025

Acknowledgements

This report has been written collaboratively, drawing on the insights and contributions of Commissioners and their organizations, as listed below. As such, the views expressed herein are not attributed to any single organization or individual. The report was authored by Stefaan Verhulst, lead technical writer and external expert, with contributions from Leona Verdadero of UNESCO. The project was chaired by UNESCO, ITU, UNDP, and the African Union Commission.

Editorial review was provided by Guilherme Canela de Souza Godoi, Cédric Wachholz, Guy Berger, Jaco du Toit, David Uribe, Tarja Turtia, David Castillo, Emma Fischer, and Lucia Bosio from UNESCO; Philippa Biggs and Nancy Sundberg from ITU; Alena Klatte and Alper Gucumengil from UNDP; Souhila Amazouz from the African Union Commission; and Begoña Otero, Adam Zable, and Roshni Singh from the Governance Lab. The production of the report was overseen by UNESCO, led by Leona Verdadero and Emma Fischer.

We would also like to thank the ITU Broadband Commission Secretariat—Nur Sulyna Abdullah, Anna Polomska, and Julia Gorlovetskaya—for their invaluable support throughout the development of this report.

Contents

List of tables, figures and boxes

Tables

Figures

Boxes

Working group members

Working group co-chairs

Ms. Audrey Azoulay, United Nations Educational, Scientific and Cultural Organization (represented by Mr. Cedric Wachholz and Ms. Leona Verdadero)

Mr Achim Steiner, United Nations Development Programme (represented by Ms Alena Klatte, and Mr Alper Gucumengil)

Ms Doreen Bogdan-Martin, International Telecommunications Union (represented by Ms Phillippa Biggs and Ms Nancy Sundberg)

HE Ms Lerato Mataboge, African Union (represented by Ms Souhila Amazouz)

Commissioners and focal points

Baroness Beeban Kidron, 5Rights Foundation (represented by Ms Marie-Eve Nadeau and Ms Nicola White)

HE Deemah Al Yahya, Digital Cooperation Organization (represented by Mr Hassan Nasser and Mr Ahmad Bhinder)

Dr. Hessa Al Jaber, Es'Hailsat

Mr Piotr Dmochowski-Lipski, EUTELSTAT IGO

Dr Qu Dongyu, Food and Agriculture Organization (represented by Mr Máximo Torero, Mr Henry Burgsteden, and Mr Cristopher Marchini)

H.E Jessica Rosenworcel, Federal Communications Commission (represented by Ms Roxanne McElvane Webber)

Mr Mats Granryd, Global System for Mobile Communications Association (represented by Mr Noriswadi Ismail, Mr Luca Elmosi, and Mr Roddy McGlynn)

Dr Rumman Chowdhury, Humane Intelligence

Mr Chuen Hong Lew, Infocomm Media Development Authority, Singapore (represented by Mr Bertrand Chew and Ms Evelyn Goh)

Ms Pamela Coke-Hamilton, International Trade Centre (represented by Mr James Howe and Mr Gilles Chappell)

Mr Lacina Koné, Smart Africa (represented by Ms Aretha Mare, Ms Thelma Efua Quaye, and Ms Denyse Ntaganda)

H.E Engineer Majed Al Mesmar, Telecommunications and Digital Government Regulatory Authority, United Arab Emirates (represented by Mr Humaid Ali Al Basti and Mr Abdul Rahman Al Marzouqi)

Ms Sima Sami Bahous, UN Women (represented by Mr Papa Seck, Ms Jessamyn Encarnacion, Ms Helene Molinier, and Ms Ramya Emandi)

Mr Filippo Grandi, United Nations High Commissioner for Refugees (represented by Mr Conor Flavin)

Mr. Ziyang Xu, Zhongxing Telecommunication Equipment (represented by Dr Zhi Cheng Qu and Mr Dao Tian)

External experts

Dr Stefaan Verhulst, The GovLab

Dr Jennifer Prendki, Quantum of Data

Dr Lucie-Aimée Kaffee, Hugging Face

Mr Bart Rosseau, Digital Flanders

Acronyms and abbreviations

AI	Artificial Intelligence
API	Application Programming Interface
ASEAN	Association of Southeast Asian Nations
AU	African Union
CARE principles	Collective Benefit, Authority to Control, Responsibility, Ethics Principles
CBDF	Cross-Border Data Flows
CDO	Chief Data Officer
DAMA-DMBOK	Data Management Body of Knowledge
DPI	Digital Public Infrastructure
DPIA	Data Privacy Impact Assessments
DPO	Data Protection Officer
DSA	Data Sharing Agreements
DSD	Digital Self-Determination
FAIR principles	Findable, Accessible, Interoperable, Reusable Principles
HRIA	Human Rights Impact Assessments
ICT	Information and Communication Technology
IT	Information Technology
ITU	International Telecommunication Union
MCC	Model Contractual Clauses
MOU	Memorandums Of Understanding
OECD	Organisation for Economic Co-Operation and Development
PIA	Privacy Impact Assessments
SDG(s)	Sustainable Development Goal(S)
SLA	Service-Level Agreements
TPM	Technical Protection Measures
UN CEB	United Nations Chief Executives Board
UNDP	United Nations Development Programme
UNESCO	United Nations Educational, Scientific and Cultural Organization
UNHCR	United Nations High Commissioner for Refugees

Introduction to the Toolkit

In an increasingly data-driven world, the ability to govern data responsibly, effectively, and for all has become a defining challenge for governments, institutions, and societies. From climate action and health systems to transportation, education, and AI development, data plays a central role in delivering public value. Yet most countries and institutions continue to face gaps in how they govern data—balancing rights, responsibilities, and reuse across fragmented systems.

This Data Governance Toolkit was developed by the Broadband Commission for Sustainable Development, through its Working Group on Data Governance, chaired by UNESCO, and co-chaired by ITU, UNDP, and the African Union. It reflects a collaborative effort among experts, policymakers, practitioners, civil society and business committed to advancing human rights-based and equitable approaches to data governance.

The Toolkit also benefitted from feedback and insights gathered through regional consultations, ensuring that it is responsive to the diverse challenges and opportunities faced by countries across different regions. These consultations have helped shape the Toolkit, making it more adaptable to local contexts while maintaining its global relevance.

It seeks to design a modular resource to support public institutions, civil society, industry and other stakeholders in designing and implementing data governance systems that are both fit-for-purpose and adaptable to local realities. Rather than prescribing a single model, it provides guiding questions, flexible frameworks, and curates actionable tools to help users navigate the full data lifecycle—from collection and storage to sharing, analyzing and use.

Governance in this context is not merely about compliance or control. It is about defining the purpose, principles, processes, people, and practices needed to ensure that data is used in ways that are trusted, valuable, equitable, and aligned with human rights. The toolkit emphasizes stewardship over ownership, recognizing that data is often co-produced, multi-use, and shaped by broader social, cultural, and political dynamics.

What's Inside

Following the "Introduction to Data Governance" this toolkit is organized around four foundational data governance components—referred to as the 4Ps of Data Governance:

- **WHY** (Purpose): seeks to help the reader identify and specify the intended goals of data and data governance and defines the public value they seek to generate.
- **HOW** (Principles): seeks to help the reader list the norms that would guide processes and practices across the data-lifecycle.
- **WHO** (People and Processes): seeks to help the reader structure and describes the roles, responsibilities, and institutional processes required to design, implement and oversee governance effectively (including the function of data stewards).
- **WHAT** (Practices and Mechanisms): seeks to help the reader outline actionable practices and mechanisms to implement decisions across the data lifecycle in alignment with the principles to meet the purpose.

The toolkit also includes:

- A self-assessment framework to help organizations evaluate their current capabilities;
- A glossary of key terms to foster shared understanding; and
- A curated list of other toolkits and frameworks for deeper engagement.

Who Is This Toolkit For?

The toolkit supports multiple types of users—each with unique priorities, responsibilities, and capacities.

The primary target group for the toolkit include:

- **Policy Leaders and Strategists.** *If you are shaping national or sectoral data strategies.* Use this toolkit to define a vision and align your country's or organization's data policies with global principles, sustainable development goals, and the needs of digital public infrastructure.
- **Chief Data Officers, Data Stewards, Data Protection Authorities, and Governance Leads.** *If you oversee how data is collected, shared, and used across your organization.* Use the toolkit to evaluate maturity levels, define principles, standardize lifecycle practices, and engage in cross-sector collaboration.

In addition, other experts who may find value in this toolkit include:

- **Legal and Compliance Professionals.** *If you ensure that data practices meet national and international laws and standards.* Use the toolkit to assess current safeguards, embed human rights protections, and align policies with data governance instruments and frameworks.
- **Technology Architects and System Designers** *If you are building platforms, APIs, or infrastructures that manage or move data.* Use the toolkit to embed governance by design, such as access controls, federated architectures, and metadata standards.
- **Multilateral and Donor Agencies** *If you are advising or funding country-level data transformation efforts.* Use the toolkit to support capacity-building initiatives, and encourage equitable, rights-based governance models.
- **Civil Society and Community Advocates.** *If you want to shape more equitable data governance.* Use this toolkit to identify points of intervention and learn about principles and practices such as Digital Self-Determination and data stewardship.

How to Engage with the Toolkit

The toolkit is modular and can be used in whole or in part:

- To develop or revise a data governance framework, begin with the four guiding questions: WHY (Purpose), HOW (Principles), WHO (Roles), and WHAT (Practices).
- To assess your organization's current status, use the Self-Assessment Tool to understand readiness and identify areas for improvement.
- To engage stakeholders, review the suggested tools to co-design policies and build legitimacy.

Limitations of the Toolkit

While this Data Governance Toolkit offers a comprehensive framework and practical tools to support responsible and effective data governance, it is important to acknowledge its limitations:

1. Contextual Adaptation Required

This toolkit is designed to be adaptable across diverse sectors, institutions, and geographies. However, its general guidance clearly cannot fully capture the legal, political, cultural, or technological nuances of every context. Users are encouraged to tailor recommendations to their specific environment – while keeping the alignment with international human rights law, and to engage local stakeholders in this process.

2. Not a Legal Instrument

The toolkit is not a substitute for legal advice or regulatory compliance. While it references examples of international principles and governance models, it does not provide authoritative interpretation of binding laws or treaties. Users should consult legal professionals and regulatory bodies to ensure alignment with applicable legal frameworks (e.g., regional and national data protection laws).

3. Evolving Landscape

The data governance field is rapidly evolving, shaped by emerging technologies (such as AI and synthetic data), new regulatory instruments, and shifting societal expectations. While this toolkit incorporates the most up-to-date practices at the time of publication, ongoing iteration will be required to stay current with future developments.

Looking Ahead

In summary, this toolkit is not intended to be prescriptive. Rather, it is a flexible resource that supports contextual adaptation, enabling each user to tailor data governance structures to their institutional, and cultural realities— taking into consideration emerging international norms and ethical imperatives.

To address these limitations, users are encouraged to:

- Treat the toolkit as a starting point for dialogue and co-creation;
- Document and share lessons from implementation to strengthen collective knowledge; and
- Participate in future updates and contribute to the growing community of practice around data governance.

Introduction to Data Governance

The twenty-first century is characterized by an accelerating process of 'datafication'—the systematic transformation of everyday practices, institutional operations, and societal systems into quantifiable data. There has been an exponential growth of data flows, driven by the proliferation of digital interactions, social media activity, and the widespread deployment of sensors and personal devices. The emergence and deployment of new technologies, like artificial intelligence (AI) further amplify this trend, both by intensifying the demand for data and by producing novel forms of data, including synthetic outputs. The development of Digital Public Infrastructure (DPI) also highlights the imperative for secure, interoperable data systems that can ensure equitable access to digital services and support large-scale, data-driven innovation.

The rapid datafication of society presents significant challenges and opportunities for public policies. Governments must adapt their frameworks to ensure data is governed securely and equitably, while also fostering innovation and protecting citizens' rights. This requires policies that address issues such as data privacy, interoperability, and ethical use of AI. As data flows become increasingly central to the functioning of digital economies and societies, effective policy frameworks are essential to ensure that data benefits all and upholds fundamental human rights.

Questions surrounding how data is governed are therefore becoming more significant. Data governance is all the more important because data, like many digital technologies, is a double-edged sword: it can drive tremendous social benefits (e.g., by improving targeted healthcare interventions and education), yet it can also cause tremendous harm (e.g., by enabling surveillance, furthering bias, and leading to privacy breaches).

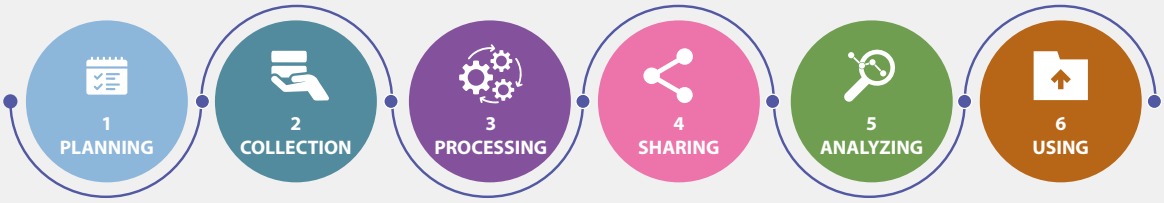
The impact of these outcomes partly depends on the effectiveness and responsiveness of data governance frameworks—the principles, processes and practices that surround how data is collected, stored, and deployed. This introduction provides a broad overview of the concept and practice of data governance.

Box 1. Understanding the Data Lifecycle

The data lifecycle refers to the various stages data goes through—from its initial planning to its ultimate use in decision-making.

While different frameworks may highlight different stages or use varying terminology, the most commonly recognized phases include:

Figure 2: Data Lifecycle



- **Planning** – Identifying data needs, intended uses, and governance requirements.
- **Collection** – Gathering data through surveys, sensors, transactions, or other means.
- **Processing** – Cleaning, validating, organizing, storing, and preserving data for use, including deletion when necessary, and ensuring proper handling throughout the data lifecycle.
- **Sharing** – Making data accessible to others for re-use, whether through platforms, APIs, or data collaboratives.
- **Analyzing** – Interpreting the data to generate insights.
- **Using** – Applying those insights to inform decisions, policies, or services.

At each stage, data governance decisions—such as who has access, how data quality is maintained, and how privacy is protected—are made. These decisions are cumulative and can significantly shape what is possible in later stages. Poor governance early on (e.g., unclear purpose or unstructured collection) can generate negative consequences, or limit the value or usability of data downstream.

1. Why Does Data Governance Matter?

There are several reasons why developing a data governance framework is critical. Among them (non-exhaustive):

- **Achieving Sustainable Development Goals (SDGs):** Good quality data is essential to monitor progress, identify gaps, and inform policies that support the United Nations' 2030 Agenda for Sustainable Development.

A wide variety of evidence now supports the role that data can play in furthering the SDGs. Without robust data governance, that potential is severely hindered.

- **More Informed Decision-Making:** Governments and organizations increasingly rely on timely, accurate, and well-managed data to make data informed decisions in a variety of sectors, including health, education, agriculture, economic development, and climate action. Effective governance can

help the integrity, accessibility, and usability of data, making it a reliable foundation for policymaking.

- **Increasing Data Collaboration:** By guiding how dispersed data is brought together, data governance supports collaboration that is more systematic, equitable, and responsible—ensuring public benefit while safeguarding key rights.
- **Mitigating Risks:** Datafication offers many benefits, but it also poses a variety of societal risks—ranging from privacy violations and security breaches to economic exclusion and political polarization. Strong data governance frameworks can help safeguard against these risks. And help maximize the positive potential of data while limiting its harms.
- **Furthering Ethical and Responsible AI:** Without high-quality and representative data and strong data governance, AI applications risk perpetuating biases, violating privacy and intellectual property, and amplifying inequalities. Data governance can thus be understood as one corner stone of ethical and responsible AI.
- **Designing Digital Public Infrastructure (DPI):** Data plays a dual role in Digital Public Infrastructure (DPI): it is both a critical enabler and, increasingly, a core component. According to the World Bank, DPI refers to "foundational systems that enable the delivery of essential society-wide functions and services," encompassing digital identity, payments, and data exchange platforms. For DPI to function effectively and deliver on its promise of inclusive, scalable, and resilient services, data should be accessible, trustworthy, and well-governed. Countries that have successfully scaled data use have done so by repeatedly reusing DPI components—such as identity systems and registries—across sectors including health, education, financial inclusion, agriculture, and land and carbon tracking. This reusability enhances efficiency, and also drives the development of new data-driven services. When DPI includes datasets or data exchange systems, it becomes a resource for public service providers, as well as private sector actors, civil society, and innovators—multiplying its value across the entire ecosystem. To unlock this potential, robust data governance is essential to ensure that these systems are interoperable, secure, and compliant with legal, ethical, and human rights standards.
- **Improving Open Government Policies:** Effective data governance frameworks also play a crucial role in open government policies. By establishing concrete rules for records management, including key datasets, and ensuring access to information in both active and passive modes, data governance becomes an essential element of any open government strategy. It ensures transparency, accountability, and the equitable distribution of public information, which fosters public trust and enhances participation in democratic processes.

2. What is Data Governance

There is no universally agreed-upon definition for data governance. The notion constitutes a broad matrix of laws, policies, norms, technologies, individuals, institutions and other mechanisms and stakeholders that includes at least two key aspects:

- How decisions about data are made;
- How data is governed throughout its lifecycle.

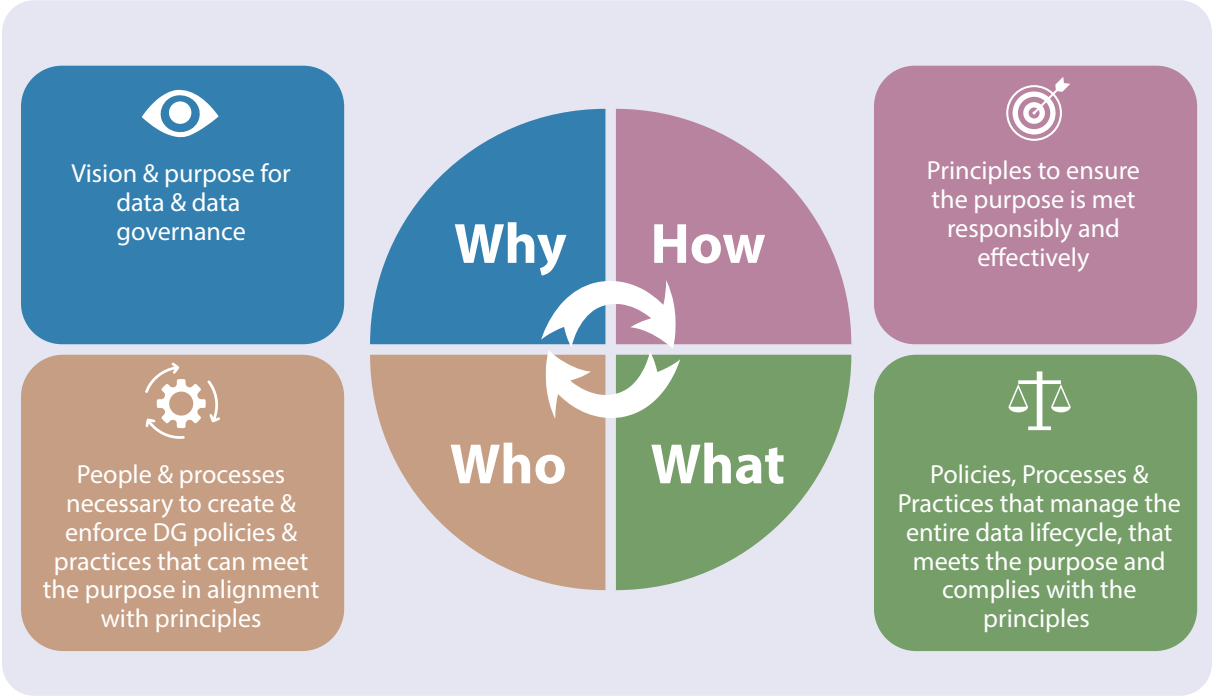
Based on this understanding, this Toolkit proposes the following working definition for data governance:

The processes, people, policies, practices and technology *that seek to govern* ***the data lifecycle*** *toward meeting the* ***purpose*** *of increasing trust, value and equity, while minimizing risk and harm in alignment with a set of* ***core principles****.*

As previously mentioned, when defining a data governance framework, four key elements and activities need to be specified:

- ***Why:*** Defining the vision and **purpose**(s) for data and data governance.
- ***How***: Specifying the **principles** that will guide and determine how decisions are made and practices implemented to meet the purpose responsibly and effectively.
- ***Who:*** Establishing **processes** and engaging the **people** necessary for creating and enforcing policies and practices that can meet the purpose(s) in alignment with the principles.
- ***What:*** Specifying and implementing the **policies, practices and technologies** that govern the different stages of the data lifecycle, in ways that meet the purpose and comply with the principles.

Figure 1. Data Governance Framework



Source: Authors.

3. Human Rights and Data Governance

This toolkit advocates taking a Human Rights-based approach to data governance. This means ensuring that data practices—across the full data lifecycle from collection to (re()use—should respect, protect, and fulfill the rights and freedoms of individuals and communities. This also entails treating data governance not just as a technical or compliance function, but as a rights-centered data governance practice.

The section below provides a few examples on how data governance can be aligned with international human rights frameworks to prevent harm and promote dignity, agency, and justice.

Privacy & Data Protection:

Conflict: Data collection and processing can help improve outcomes in fields as diverse as educational performance or medical diagnosis and treatment. However, data processing could potentially infringe on the individual's right to privacy, especially when involving sensitive personal information or where security safeguards are not regularly reviewed and updated. *Human Rights- Based Response:*

- Adopt data minimization, privacy by-design, encryption, and secure storage practices.
- Ensure informed, meaningful consent and transparency in how data is used.
- Enable individuals to access, correct, or delete their data.
- Limit data retention to what is necessary, proportionate and justifiable under fundamental rights standards.

Discrimination & Bias:

Conflict: Biased data and algorithmic systems can reinforce systemic discrimination in data analysis and decision-making, resulting in discriminatory outcomes (either favoring one population or producing adverse results for another, or both) and denying equal treatment to all communities. *Human Rights- Based Response:*

- Audit data and models to identify, explain and mitigate bias and discriminatory outcomes.
- Engage diverse voices in data governance processes.
- Ensure datasets reflect and represent diverse populations and are contextually appropriate.
- Monitor systems continuously to identify and correct disparate impacts.

Surveillance & Mass Data Collection

Conflict: Surveillance of certain locations (e.g. cash machines, airports, public facilities) or events (e.g. football matches, elections or demonstrations) can help deter violence or improve security. However, unchecked surveillance without proper limitations can undermine privacy, free expression, self-determination and the right to peaceful assembly. *Human Rights- Based Response:*

- Enact and enforce strong legal safeguards, oversight, and due process.
- Limit surveillance and data collection to purposes that are lawful, specific, necessary and proportionate.
- Increase transparency around surveillance programs and technologies, making them accessible and open to challenge.
- Provide independent accountability mechanisms to review abuses.

Freedom of Expression & Access to Information

Conflict: Data governance, including data access controls, can lead to censorship or restrict access to vital information. *Human Rights- Based Response:*

- Establish fair, transparent, and accountable data access policies.
- Promote open access to public-interest information and data.
- Ensure equitable access to data.
- Strengthen digital literacy and civic engagement.
- Follow the 9 principles of a Freedom of Information Regime

4. Accountability & Redress

Conflict: A strong, fair and effective data governance framework can help protect individuals and their data. Opaque or unaccountable data governance can leave individuals without recourse for violations. *Human Rights- Based Response:*

- Create clear, timely and accessible grievance mechanisms for data-related rights abuses.
- Guarantee access to legal remedies, including administrative, judicial or ombudsperson pathways, with enforceable reparations.
- Promote transparency in data governance decisions and risk assessments.
- Include and ensure public participation in data governance frameworks.

Data Governance Checklist and Self-Assessment

The Data Governance Self-Assessment is designed to help governments, public institutions, and organizations evaluate the current state of their data governance systems and identify opportunities for improvement. It provides a structured way to move from reflection to action, based on the 4Ps framework: Purpose, Principles, People, and Practices.

STEPS FOR USING THE SELF-ASSESSMENT

Below are step-by-step instructions and key considerations for completing the tool effectively.

Step 1: Assemble a Cross-Functional Team

Assemble a cross-functional group of internal and external stakeholders

If possible, nominate a facilitator or data steward to guide the process and consolidate inputs.

Step 2: Review Each Domain and Dimension

The assessment is divided into six major domains:

1. Vision & Purpose
2. Principles
3. Processes & Roles
4. Practices & Policies
5. Emerging Trends & AI Governance
6. Capacity & Literacy

For each domain, you'll find a set of Yes/No questions designed to assess whether key elements are in place.

Note: This is not a compliance exercise but a tool for learning and prioritization.

Step 3: Determine Priority Level

Based on your strategic goals and constraints, indicate the priority for improving each dimension:

1. High: Immediate action needed; foundational to other areas
2. Medium: Important but not urgent.
3. Low: Lower relevance or already functioning well.

Step 4: Analyze Results and Plan Next Steps

Identify critical gaps

Use results to inform your data governance strategy, capacity-building plans, or policy revisions.

Share findings internally or with partners to foster alignment and coordinate efforts.

Use periodically (e.g., annually) to track progress over time and ensure continuous improvement.

The Data Governance Toolkit

Readers can refer to the other sections of the Data Governance Toolkit for targeted support:

- Use the WHY (*Purpose)* section to refine strategy;
- Revisit HOW (*Principles)* to align with evolving norms;
- Consult WHO (*Processes and People)* for institutional design and stewardship roles;
- Apply WHAT (*Practices)* to consider mechanisms across the data lifecycle.

The Checklist and Self-Assessment Tool

WHY – Purpose:

Dimension	Question	Yes	No	Priority Level	Comments
Vision & Purpose	Is there a national data strategy or vision that explicitly links data (re)use to priorities (e.g., climate, health, digital transformation)?	☐	☐		
	Do agencies or sectors have their own formal data strategies or plans?	☐	☐		
	Are data-driven objectives or initiatives referenced in official policy documents (e.g., agency strategies, policy briefs, implementation plans) in at least some domains?	☐	☐		
	Are public value use cases or priority areas for data use identified (e.g., in sectors like health, education, or crisis response)?	☐	☐		
	Are these use cases or priority areas co-developed with key stakeholders, including citizens, civil society, or frontline service providers?	☐	☐		
	Is there a formal process or mechanism for regularly reviewing and updating the national data strategy based on evolving needs or stakeholder input?	☐	☐		

HOW – Principles:

Dimension	Question	Yes	No	Priority Level	Comments
Principles (Documentation)	Are data-related principles (e.g., transparency, accountability, fairness, participation) officially documented or publicly endorsed at the political or organizational level?	☐	☐		
	Are separate technical or operational data governance principles (e.g., data quality, stewardship, access control) formally defined or adopted at the agency/department level?	☐	☐		
Principles (Implementation & Alignment)	Are these principles understood by key personnel and embedded into operational practices across data lifecycle stages (e.g., collection, processing, sharing, reuse, storage)?	☐	☐		
	Are the adopted principles aligned with key international frameworks such as **FAIR** (Findable, Accessible, Interoperable, and Reusable), **CARE** (Collective Benefit, Authority to Control, Responsibility, and Ethics), and **FIPS** (Federal Information Processing Standards), ensuring consistency with global best practices in data governance.	☐	☐		
	Are these principles enforceable through contracts, MOUs, or data-sharing agreements with external partners?	☐	☐		

WHO – People & Roles:

Dimension	Question	Yes	No	Priority Level	Comments
People & Roles	Is there a designated Chief Data Officer (CDO), Chief Data Protection Officer (DPO), Chief Data Steward, or equivalent senior data leader with a formal mandate and adequate resources?	☐	☐		
	Is there an independent data protection regulator or authority in place to oversee data privacy and protection practices, ensuring compliance with national and international data protection laws?	☐	☐		
	Are the responsibilities of key data governance roles (e.g., CDO, DPO, AI ethics board, data steward) clearly defined and documented?	☐	☐		
	Are formal or informal roles or structures in place for data stewardship, privacy oversight, and ethical data use within public sector institutions (e.g., data stewards, privacy officers, ethics boards)? Please use the comments section to explain.	☐	☐		
	Are there coordinating mechanisms in place to align and determine common policies and practices?	☐	☐		
	Is there any collaboration (structured or unstructured) on data governance between government agencies (e.g., inter-agency data committees, joint strategies, shared data platforms)? Please use the comments section to explain.	☐	☐		
	Is there any coordination (formal or informal) on data governance with actors outside the government (e.g., academia, civil society, private sector)? Please use the comments section to explain.	☐	☐		

WHAT – Practices (Policies & Governance)

Dimension	Question	Yes	No	Priority Level	Comments
Policies & Governance	Are formal data governance policies covering data collection, processing, sharing, and reuse adopted and publicly available?	☐	☐		
	Were these policies developed using a participatory approach?	☐	☐		
	Do all departments apply these policies consistently?	☐	☐		
	Are these policies aligned with ethical and legal standards (e.g. national and regional laws)?	☐	☐		
	Are policies integrated with Freedom of Information legislation, ensuring that data accessibility and disclosure align with public access to information requirements?	☐	☐		
	Are risk and impact assessments mandatory before data is used or shared, particularly for sensitive data or high-risk datasets (health, finance, demographics)?	☐	☐		
	Do requirements vary between statistical offices, IT units or AI teams?	☐	☐		

WHAT – Practices (Data Lifecycle)

Dimension	Question	Yes	No	Priority Level	Comments
Data Collection & Storage	Are data collection processes aligned with purpose limitation and minimization principles?	☐	☐		
	Is data expiry (end of life) or retention specified?	☐	☐		
	Is consent (including dynamic consent) obtained where necessary?	☐	☐		
	Do data collection methods consider the representation of marginalized or underserved groups?	☐	☐		
	Are these practices standard across agencies, or specific to certain programs or units?	☐	☐		
Data quality, Management & Metadata	Are data quality controls in place (e.g., accuracy, completeness, timeliness)?	☐	☐		
	Are data catalogs used, and are metadata standards (e.g., DCAT, FAIR) applied?	☐	☐		
	Is data lineage practices in place as to ensure traceability and increase accountability?	☐	☐		
	Are metadata and quality practices aligned across departments or led by specific units (e.g., national statistics office)?	☐	☐		
Data Sharing	Are there policies and technical tools to enable secure and ethical data sharing and reuse (e.g., data sharing agreements, APIs, federated access models)?	☐	☐		
	Is data discoverable through catalogs, portals, or registries?	☐	☐		
	Are data-sharing tools and portals shared across agencies or developed in silos?	☐	☐		

Data Security & Compliance	Are there legal and technical safeguards for data security and privacy?	☐	☐		
	Are data access and sharing aligned with national or regional legal frameworks?	☐	☐		
	Are foundational digital public infrastructure (DPI) elements like registries and IDs securely managed?	☐	☐		
Data Use	Are data analytics (incl. visualization) used in policymaking?	☐	☐		
	Are mechanisms in place for responsible reuse of private-sector data for public interest purposes?	☐	☐		
	Is data used for crisis response (e.g., pandemic, natural disasters)? If 'yes', which departments or domains lead in data-driven use cases?	☐	☐		
Inclusivity, Equity, Self Determination	Are marginalized groups represented in data governance processes?	☐	☐		
	Are there appropriate frameworks supporting Indigenous or community control and governance of data where relevant?	☐	☐		
	Is equitable access and use of data a formal consideration in governance or practice?	☐	☐		

WHAT – Practices (AI & Emerging Trends):

Dimension	Question	Yes	No	Priority Level	Comments
Emerging Trends and Technologies	Are there AI governance frameworks ensuring datasets are AI-ready (e.g., unbiased, representative, high-quality)?	☐	☐		
	Are there ethical guidelines for AI development using public data?	☐	☐		
	Are privacy enhancing technologies used to minimize risks?	☐	☐		

WHAT – Capacity and Literacy:

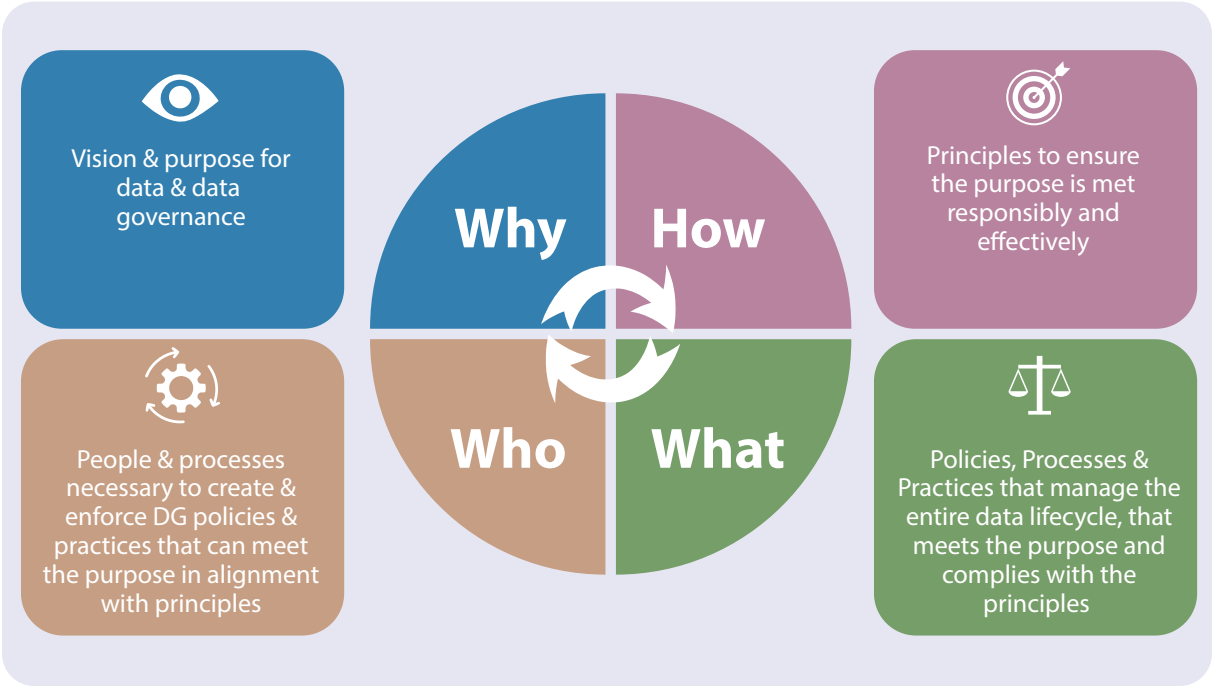
Dimension	Question	Yes	No	Priority Level	Comments
Capacity and Literacy	Do actors at multiple levels demonstrate a basic understanding of data governance concepts?	☐	☐		
	Are there regular training programs on data governance, privacy, AI ethics, and compliance?	☐	☐		
	Are there formal or informal communities of practice working on data governance (e.g., learning circles, working groups)?	☐	☐		
	Are there mechanisms to collaborate with universities, civil society, or international organizations to strengthen institutional knowledge?	☐	☐		
	Is there a dedicated budget or long-term investment plan for capacity-building in data governance and stewardship?	☐	☐		
	Are there communities of practice and knowledge-sharing platforms for public servants?	☐	☐		
	Are civil servants equipped to engage in data collaboration and partnerships (incl. cross-sector)?	☐	☐		

Building blocks for developing a data governance framework

While different sectors and jurisdictions may require tailored data governance approaches, the following section seeks to provide building blocks that can be leveraged and adapted to guide the development of new or refinement of existing data governance strategies. As explained earlier, data governance can be deconstructed along the following four questions (the 4Ps of data governance) and actions that need to be specified:

- ***Why:*** Defining the vision and **purpose** for data and data governance.
- ***How:*** Specifying the **principles** that will guide and determine how decisions are made and practices implemented to meet the purpose responsibly and effectively.
- ***Who:*** Establishing **processes** and identifying **people** necessary to create and enforce policies and practices that can meet the purpose in alignment with the principles.
- ***What:*** Specifying and implementing the **policies, practices and technologies** that govern the different stages of the data lifecycle, in ways that meet the purpose and comply with the principles.

Figure 1.



Other Data Governance Toolkits

Several other data governance toolkits already exist; a curated selection is presented below, with an indication of their intended audiences. This Toolkit does not intend to duplicate them, but to complement them. The value of this Toolkit lies in its comprehensive approach, beginning with the purpose, principles, and processes of data governance. Unlike many existing frameworks that focus directly on practices, the Toolkit emphasizes the foundational elements, ensuring that data governance is grounded in strategic and ethical considerations. This structure enables more effective, adaptable, and aligned governance across diverse contexts.

Table 1. Data Governance Toolkits Mapping

Toolkit	Audience
Data Governance Toolkit (NSW Government Agencies)	Public Sector
Data Innovation Toolkit (European Commission Digital Innovation Lab)	Public Sector
OECD Data Governance	Public Sector
Data to Policy Navigator (UNDP)	Public Sector
Data policy Framework (African Union)	Public Sector
Data Management Framework (ASEAN)	Public Sector
Navigating Data Governance (ITU)	Regulators
The Data PlayBook (IFRC and Solferimo Academy)	Humanitarian Sector
Data Responsibility Journey (The GovLab)	Public and Private Sector
Data Governance and Management Toolkit (SGIG DSC Members)	Self-Governing Indigenous Governments
Data Governance Workbook (Digital Civil Society Lab)	Non-Profit Sector

I. WHY

Defining the vision and purpose for data and data governance

Yurchanka Siarhei/Shutterstock.com

Core Objective	Why Principles Matter
To establish and evaluate a clear, actionable vision and purpose for data use and governance within an organization or government.	A clear vision and purpose(s) are the bedrock of effective data governance. They articulate why an organization or country collects, processes, and uses data, reflecting its core values and strategic priorities. Purposes are not monolithic; they can span a wide range of objectives, from maximizing data's economic and social value and fostering innovation to establishing equity, supporting specific policy goals (like health outcomes or environmental protection), or ensuring responsible AI development. Defining the purposes prompts crucial questions about who sets these values and ensures they align with broader societal goals and international human rights law.

Assessing the Vision and Purpose

To determine whether a meaningful vision and purpose(s) for data and data governance are in place and effectively implemented, specific elements should be assessed. The following questions probe the existence, nature, and application of this strategic direction.

1. Strategic Alignment and Intent: Connecting Vision to Strategy

- **Question: Is there a national data strategy or vision that explicitly links data (re)use to policy priorities (e.g., climate, health, digital transformation)?**
  - **Rationale**: A national strategy is the primary instrument for translating a high-level vision into a coherent plan. It signals commitment and should articulate which specific purposes for data are being prioritized at the national level. Does the strategy focus primarily on maximizing data utility and value for economic growth? Is there a strong emphasis on fostering innovation and long-term development (e.g., advancing the SDGs)? Does it explicitly aim for establishing equity and self-determination in the data ecosystem? Or does it concentrate on supporting specific policy objectives like open data for transparency, enabling cross-border data sharing, or mobilizing data for crisis response? A clear strategy would connect the overarching vision to these selected purposes and align data governance with pressing policy challenges.
- **Question: Do individual agencies or sectors have their own formal data strategies or plans?**
  - **Rationale**: This assesses whether the national vision and its prioritized purposes are cascaded down and translated into actionable plans within specific operational contexts. Agency-level strategies operationalize the broader goals, defining how that specific entity will contribute. For example, a health agency's strategy might detail how it will use data to advance specific SDGs related to health (Supporting Specific Policy Objectives) while implementing governance frameworks to protect vulnerable groups' data (Establishing Equity).

- **Question: Are data-driven objectives or initiatives referenced in official policy documents (e.g., agency strategies, policy briefs, implementation plans) in at least some domains?**
    - **Rationale**: This question verifies whether the strategy (and the purposes it embodies) is integrated into the organization's core work, ensuring data is treated as a critical asset for achieving outcomes, rather than being confined to a standalone document.

2. Defining Value and Prioritizing Use: Translating Purpose into Action

- **Question: Are public value use cases or priority areas for data use identified (e.g., in sectors such as health, education, or crisis response)?**
    - **Rationale**: A clear vision and strategy should translate into concrete applications that deliver tangible value. Identifying priority use cases demonstrates how the chosen purposes are being pursued. These use cases should directly reflect strategic priorities. For instance:
        - Maximizing Utility/Value: Use cases focused on economic insights, optimizing services, or generating new revenue streams.
        - Fostering Innovation/SDGs: Initiatives promoting open data ecosystems, data-driven research for sustainable development goals (health, environment, etc.) or skills development.
        - Establishing Equity: Use cases designed to address disparities, improve access for marginalized groups, or protect data subjects' rights.
        - Supporting Specific Policy Objectives: Use cases targeting enhanced government transparency (Open Data), improved disaster management (Crisis Response), fostering international research (Cross-border Sharing), ensuring fair AI deployment (Harnessing AI), or building interoperable systems (Aligning with DPI). This assessment verifies that the strategic intent leads to focused, impactful data utilization.

3. Inclusivity and Stakeholder Engagement:

- **Question: Are these use cases or priority areas co-developed with key stakeholders, including citizens, civil society, or frontline service providers?**
    - **Rationale**: Data governance does not exist in a vacuum. Co-development ensures that the defined purpose and priorities reflect the actual needs and values of those affected by data use. It promotes equity and self-determination by including diverse perspectives, especially from marginalized or vulnerable groups, helping to establish trust and ensure data initiatives are genuinely beneficial and minimize harm. This aligns with the use of data in a just way.

4. Adaptability and Continuous Improvement:

- **Question: Is there a formal process or mechanism for regularly reviewing and updating the national data strategy, with input from a diverse range of stakeholders, to ensure it remains aligned with evolving national needs and priorities?**
    - **Rationale**: The data landscape is dynamic. A formal review process ensures the vision and purpose remain relevant and effective amidst changing technologies, societal expectations, and policy goals. It allows for course correction and demonstrates a commitment to ongoing improvement and accountability.

Table 2. Checklist: Potential Purposes for Data and Data Governance

Please check all purposes that apply to your data governance initiative:

Maximizing Data Utility and Value / Minimizing Harms and Costs	
Unlock data for new insights and improved decision-making.	☐
Create tangible value for the organization and/or society through data use.	☐
Minimize potential harms, costs, and unintended negative consequences associated with data collection and use.	☐
Fostering Innovation and Sustainable Development	
Stimulate innovation and entrepreneurship using data-driven approaches.	☐
Promote and enhance economic opportunities through data.	☐
Advance Sustainable Development Goals (SDGs) using data for monitoring and implementation (e.g., in health, education, infrastructure).	☐
Foster a data-driven culture and enhance data skills within the organization/society.	☐
Establishing Equity and Digital Self-Determination	
Promote equitable access to data and ensure benefits are shared fairly.	☐
Ensure groups in situation of marginalization and vulnerability, including children, are actively included and protected in data ecosystems.	☐
Uphold principles of data self-determination for individuals and communities.	☐
Implement specific governance frameworks to protect the data of vulnerable populations.	☐
Supporting Specific Policy or Operational Objectives	
Enhance transparency, accountability, and citizen engagement (e.g., through Open Data initiatives).	☐
Support national SDG priorities such as accommodating the impact of climate change or advancing education for all.	☐
Facilitate safe and effective cross-border data sharing and international collaboration while ensuring legal compliance.	☐
Mobilize data effectively for crisis preparedness, risk management, response, and recovery efforts.	☐
Support the responsible development and deployment of Artificial Intelligence (AI), addressing bias, fairness, and ethical concerns.	☐
Align with or implement Digital Public Infrastructure (DPI) principles to enable secure, private, interoperable, and seamless data exchange.	☐

Methods

Data governance is about ensuring data is governed effectively to serve a country or organization's goals and values. Defining the purpose of data governance is the critical first step, as it dictates the scope, priorities, and ultimate success of any governance initiative. The methods listed below each play a distinct yet often interconnected role in this definitional process, and are particularly strengthened when viewed through a multistakeholder lens.

Understanding Needs and Perspectives:

- Stakeholder Mapping: This is foundational. Before defining purpose, it is necessary to know who is impacted by data governance nd who has a vested interest in its outcome. Mapping identifies key individuals, teams, and external parties (like regulators or partners) and their current relationship with data. Understanding their pain points, needs, and expectations regarding data quality, access, security, and privacy is crucial for shaping a relevant and valuable governance purpose. Example: Stakeholder Mapping and Stakeholder Engagement

Setting Vision and Goals:

- Vision and Mission Workshops: These collaborative sessions bring stakeholders together to brainstorm and articulate a shared future state for data within the organization. By defining a compelling vision (the ideal future) and a clear mission (the organization's role in achieving that future through data) through a multistakeholder process, these workshops can directly contribute to defining the why of data governance. The purpose should align with and support the achievement of this broader data vision and mission. Example: Vision and Mission Workshops and Workshops and Consultation Sessions
- Theory of Change Framework: This framework helps to logically map out the causal pathways from the activities of data governance to the desired long-term outcomes and impact. By articulating the assumptions and steps required to achieve governance goals, it clarifies how data governance is expected to drive change and achieve its purpose. This process helps to refine the purpose by ensuring it is realistic and has a clear path to impact. Example: Theory of Change Framework
- Purpose Canvas: This is a dedicated tool specifically designed for clarifying purpose. A purpose canvas typically prompts reflection on elements like the core problem being solved, the target beneficiaries, the value proposition of the initiative, and the key activities. Applying this to data governance forces a structured articulation of why governance is necessary, who it benefits, and what key functions it will perform. Example: Purpose Canvas

Ensuring Alignment and Effectiveness:

- Benchmarking and Best Practices Review: Looking at how other successful organizations have approached data governance can provide valuable insights into common challenges, effective strategies, and potential purpose statements. While not a direct method for defining an organization's unique purpose, it helps in understanding what is possible, identifying potential areas of focus, and ensuring the defined purpose is comprehensive and addresses recognized best practices in the field. Example: Benchmarking and Best Practices Review
- Regulatory and Compliance Review: Data governance is often driven by regulatory requirements. A thorough review of relevant laws, including international obligations, and regulations is essential to ensure the purpose of data governance explicitly includes meeting these obligations. Compliance is not just a task; it is often a fundamental reason for implementing governance, and the purpose

should reflect this. Example: Regulatory and Compliance Review and Policy Review and Benchmarking

- Scenario Planning: This involves exploring different potential future states or challenges related to data (e.g., a data breach, a new technology adoption, a regulatory change). By considering how data governance would need to function in these scenarios, organizations can ensure their defined purpose is robust, adaptable, and forward-looking. This helps to define a purpose that is not just relevant today but will remain so in the face of future uncertainty. Example: Scenario Planning
- Balanced Scorecard: While primarily a performance management tool used after the purpose is defined, the process of developing a Balanced Scorecard for data governance can retrospectively inform and refine the understanding of its purpose. By identifying the key perspectives (e.g., financial, customer, internal processes, learning and growth) and objectives that data governance contributes to, it reinforces why governance is important and how its success will be measured in relation to organizational goals. This reinforces the link between the governance purpose and broader strategic objectives. Example: Balanced Scorecard

Resources for further reading:

- ASEAN Data Management Framework
- UNDG Big Data Guidance Note
- UNDP Data Futures Exchange (DFx)
- UNESCO Open Data Guidelines
- African Union Resolution on Promoting and Harnessing Data Access as a Tool for Advancing Human Rights and Sustainable Development in The Digital Age.

National Case studies

- UNESCO. Artificial intelligence needs assessment survey in Africa
- Carnegie Endowment for International Peace's Data Governance: Asian Alternatives: How India and Korea Are Creating New Models and Policies
- UNESCO. Kenya: artificial intelligence readiness assessment report
- UNESCO. Chile: artificial intelligence readiness assessment report
- UNESCO. (2025). South Africa: Artificial Intelligence Readiness Assessment Report.
- UNESCO. (2025). Mozambique: Artificial Intelligence Readiness Assessment Report.
- UNESCO. (2025). Indonesia: Artificial Intelligence Readiness Assessment Report.

II. HOW

Defining Principles for Data Governance

TippaPatt/Shutterstock.com

Core Objective

Define and uphold a clear set of guiding principles that inform all processes, decisions, and practices within the data governance framework, ensuring consistency, accountability, and alignment with ethical, legal, and societal values.

Why Principles Matter

Principles serve as the foundation of a data governance framework, shaping all activities, decisions, and processes to ensure consistency, accountability and compliance with agreed standards. When aligned with globally recognized frameworks, (such as human rights, data provenance and interoperability standards or ethical AI guidelines), principles can help ensure that data practices remain transparent, fair and adaptable. Embedding these principles into organizational governance structures also facilitates data collaboratives, fosters trust among stakeholders and may support cross-border data collaboration while safeguarding privacy and security.

Taxonomy of Principles

Different categories of principles guide various building blocks the data governance, highlighting their interconnected nature and often overlapping. The following provides an overview of the most common principles used to shape how data is governed across processes, decision-making, and data handling.

1. Principles for Processes

These principles influence how governance processes should be conducted, seeking to ensure that decisions are made in a fair, transparent, and equitable manner. They include:

- **Transparency:** Ensuring that governance processes, methodologies, and rationales are open, understandable, and accessible to relevant stakeholders. This principle is essential for upholding the human right to seek, receive, and impart information, enabling public scrutiny and fostering accountability.
- **Accountability:** Establishing clear responsibility for governance outcomes and ensuring mechanisms exist to hold individuals and organizations answerable for their actions related to data, particularly when those actions impact human rights.
- **People-Centered:** Prioritizing the needs, rights, well-being, and interests of individuals and communities throughout the governance process. This principle explicitly places human dignity and fundamental freedoms at the core of data governance design and implementation.
- **Fairness:** Seeking to ensure equal treatment, avoid arbitrary practices, and proactively mitigate bias within governance procedures. This principle directly reflects the human right to equality and non-discrimination, ensuring that governance itself does not perpetuate or create unjust disparities.
- **Participation:** Actively involving relevant stakeholders, including affected individuals and groups in situation of marginalization, in the design, implementation, and evolution of governance processes to the extent practicable and appropriate. This aligns with rights related to freedom of association, peaceful assembly, and

participation in public affairs, recognizing that those impacted by data governance should have a voice.

- **Lawfulness:** Ensuring that all governance processes and activities fully comply with applicable legal frameworks, including international law and regulations, is a fundamental principle under human rights law. This principle requires that any interference with rights must be prescribed by law (principle of legality), that the law itself must be compatible with human rights standards (principle of legitimacy), and that the provisions of the law are necessary and proportional to achieve the intended objective. These three elements are known as the three-part test, ensure that legal actions respect the fundamental rights of individuals while maintaining a lawful and balanced approach to governance.
- **Inclusion:** Ensuring that governance processes are designed to represent and include the perspectives and needs of all relevant groups, with particular attention to 'data invisible' communities, Indigenous peoples (relevant to CARE principles, see below), and stakeholders concerned about the human impacts of datafication. This principle is vital for realizing the human right to non-discrimination and ensuring that data governance serves and protects everyone equitably.

2. Principles for Decisions

These principles guide the decisions made in the governance framework, shaping how policies are defined and implemented. They include:

- **Transparency:** Clearly communicating the rationale, criteria, and impact behind governance decisions to affected parties. Transparency in decision-making is crucial for enabling individuals to understand how decisions affecting their data (and potentially their rights) are made, supporting the right to information and facilitating challenges when necessary.
- **Proportionality:** Ensuring that governance decisions, policies, and any restrictions on data use are appropriate to the scale of the data activity, the potential risks, and the intended benefits, while minimizing potential negative impacts on human rights. This principle is a cornerstone in human rights law when balancing competing rights or legitimate aims; any limitation on a right should be necessary and proportionate to the objective pursued.
- **Defined Purpose:** Ensuring that all governance decisions are made with a clear, legitimate, and specific purpose in mind, and that data use aligns with the purpose. This principle, fundamental to data protection and privacy rights, prevents arbitrary or mission-creep uses of data that could infringe on rights and freedoms.
- **Accountability:** Ensuring that decision-makers are clearly identified and held answerable for the outcomes and impacts of the policies and choices they make, especially concerning potential human rights implications. This reinforces the right to an effective remedy and encourages responsible decision-making.
- **People-Centered:** Putting individuals' needs, rights (including privacy, autonomy, and non-discrimination), and well-being at the forefront when making decisions about data collection, use, and sharing. This principle ensures that decision-making prioritizes human impacts over purely technical or organizational convenience.
- **Fairness:** Ensuring that governance decisions result in just and equitable treatment for individuals and groups, proactively working to identify and eliminate potential sources of bias in decision-making processes or outcomes.

This directly supports the human right to equality and non-discrimination.

- **Protection from Harm and Non-Discrimination:** Explicitly designing decisions and policies to safeguard against potential negative impacts on individuals or groups and actively prevent discriminatory uses of data, biased algorithmic outcomes, or other harms. This is a core human rights principle – the obligation not to cause harm and to ensure equal treatment and opportunities for all.
- **Participation:** Encouraging and incorporating diverse input and perspectives in the decision-making processes, particularly when decisions impact different stakeholder groups, including vulnerable or marginalized populations. This strengthens the legitimacy of decisions and helps ensure they respect the rights and realities of those affected.

3. Principles for Data Handling

These principles guide how data is processed, stored, and shared, ensuring responsible and secure handling. They include:

- **Confidentiality and Security:** Implementing robust technical and organizational measures to protect sensitive and confidential data from unauthorized access, disclosure, alteration, or destruction. Adequate security is not just a technical requirement; it is essential for protecting fundamental freedoms, particularly the right to privacy and potentially others like the right to security of a person, as data breaches can lead to significant human rights violations.
- **Proportionality:** Ensuring that the methods and extent of data collection, processing, storage, and sharing are appropriate and limited to what is strictly necessary and least intrusive for the defined, legitimate purpose. Where personal data is concerned, this principle is critical for upholding the right to privacy, ensuring that data handling activities do not unduly infringe upon individuals' personal spheres beyond what is justifiable and necessary.
- **Data and Information Accessibility and Portability:** Striving to make data accessible and usable where appropriate (e.g., open data principles), while balancing the needs for sharing and discovery with requirements for privacy, security, intellectual property, and user control (e.g., individuals' right to access their own data and potentially move it). This principle supports various rights, including the right to seek and receive information (access to information), the right to privacy (freedom from intrusion and control over one's data), and potentially, in some countries, freedom of expression and association in the digital environment.
- **Protection of Privacy:** Ensuring that personal data is collected, used, and managed in a manner that respects individual privacy rights and complies with relevant privacy laws and regulations. The right to privacy is a fundamental human right in many countries, enshrined in various national and international instruments, and this principle makes its protection central to all data handling activities.
- **Lawfulness:** Ensuring that all data handling activities comply with applicable legal frameworks, including data protection regulations, sectoral laws, and contractual obligations. Lawful data handling means activities have a basis in law, and critically, that these laws themselves are compliant with international human rights standards.
- **Informed Consent:** Where consent is the legal basis for processing, ensuring individuals are provided with clear, understandable information about how their data will be used and freely give

or withdraw their agreement without coercion. This principle is deeply grounded in human rights concepts of autonomy, dignity, and the right to control one's personal information.

- **Data and Metadata Quality:** Committing to maintaining high standards for the accuracy, completeness, consistency, timeliness, and relevance of data, including inferred data. Ensuring metadata is structured, comprehensive, and machine-readable facilitates data discoverability, understanding, and responsible reuse. Poor data quality can lead to unfair or discriminatory outcomes and potentially violate human rights (e.g., incorrect data impacting access to social services, or leading to biased decisions by automated systems).
- **Interoperability and Standardization:** Adhering to recognized data formats, structures, protocols, and standards to promote seamless data exchange and integration within and across systems, organizations, and jurisdictions. While primarily technical, this principle can support human rights by facilitating access to public information, enabling data portability to support individual rights, and promoting transparency through standardized data flows, provided it is implemented with privacy and security safeguards.

Box 2. Emerging Principle: What is Digital Self-Determination (DSD)?

Digital Self-Determination (DSD) is a novel principle to data governance that shifts control from centralized institutions to individuals and communities, allowing them to actively participate in how their data is used throughout its lifecycle. Unlike traditional models, which rely on one-time consent or state-centric governance, DSD introduces continuous, adaptive control and focuses on participatory decision-making based on a social license. Key principles include:

- **Dynamic, Ongoing Control**: Individuals and communities can adapt how their data is used as circumstances change, leveraging a social license.
- **Active Agency Across the Data Lifecycle**: Individuals participate not only in the collection of data but also in its processing, sharing, and deletion, fostering a more participatory governance structure.
- **Transparency and Trust**: DSD requires ongoing transparency, ensuring data subjects are continually informed about how their data is being processed, stored, and shared.
- **Ethical and Collective Focus**: DSD transcends the ownership and commodification models by embedding ethical, social, and communal dimensions into data governance. It emphasizes collective governance, allowing communities to co-manage data in alignment with individual and group rights.
- **Decentralized, Democratic Governance**: DSD decentralizes authority, encouraging shared governance among individuals, communities, and organizations rather than concentrating power in institutions.

Source: International Network for Digital Self-Determination.

Table 3. Checklist: Principles for Data and Data Governance

This checklist provides a way to determine and evaluate the presence of principles across three categories:

1. **Principles for Processes** – how governance decisions are made
2. **Principles for Decisions** – what informs those decisions
3. **Principles for Data Handling** – how data is managed in practice

Principles for Processes	
These principles can help ensure that governance activities are conducted in a fair, equitable, and transparent way.	
Transparency – Governance processes are open and understandable.	☐
Accountability – Actors are responsible for outcomes of governance processes.	☐
People-Centeredness – The needs and rights of individuals are prioritized.	☐
Fairness – Equal and unbiased treatment in procedures.	☐
Participation – Stakeholders are meaningfully involved in governance.	☐
Lawfulness – All actions comply with relevant laws and regulations.	☐
Inclusion – Marginalized groups and those affected by datafication and the increasing digitization of society are represented.	☐
Principles for Decisions	
These principles can help guide how decisions in data governance are defined and implemented.	
Transparency – The rationale behind decisions is clearly communicated.	☐
Proportionality – Decisions are appropriate to their context and impact.	☐
Defined Purpose – Decisions are guided by clear and specific objectives.	☐
Accountability – Decision-makers are answerable for their choices.	☐
People-Centeredness – Individuals' needs and rights are front and center.	☐
Fairness – Decisions are just and equitable. Protection from Harm and Non-Discrimination – Risks are mitigated, and bias is avoided.	☐
Participation – Diverse perspectives are integrated into decision-making.	☐

Principles for Data Handling	
These principles can help ensure that data is managed responsibly, securely, and in line with legal rights and expectations.	☐
Confidentiality and Security – Sensitive data is protected from unauthorized access.	☐
Proportionality – Data practices are aligned with the purpose and necessity.	☐
Accessibility and Portability – Data is available and portable under appropriate conditions.	☐
Protection of Privacy – Personal data is safeguarded according to privacy laws.	☐
Lawfulness – Data handling complies with all applicable legal standards.	☐
Informed Consent – Data subjects are aware of and agree to data use.	☐
Data and Metadata Quality – Data is accurate, reliable, and well-documented for reuse.	☐
Interoperability and Standardization – Data adheres to shared standards for easier exchange and integration.	☐

Existing and Specific Principles

Some existing examples of frameworks that can help inform data governance frameworks include:

- CARE Principles: Centered on the rights of Indigenous peoples, ensuring data governance promotes Collective Benefit, Authority to Control, Responsibility, and Ethics.
- Digital Cooperation Organization Privacy Principles: aim to provide common foundation on data privacy to the DCO Member States, and would form basis of the DCO Interoperability Mechanism for Cross-Border Data Flows.
- FIPPS (Fair Information Practice Principles): Core principles like transparency, purpose specification, accountability, data quality, and accessibility help ensure fair and responsible data handling.
- FAIR Principles for Scientific Data: Focused on making data Findable, Accessible, Interoperable, and Reusable, primarily for research data.
- FARR Principles: Fair Principles in Machine Learning (ML), AI Readiness & Reproducibility.
- IASC, Operational Guidance on Data Responsibility in Humanitarian Action.
- OECD Good Practice Principles for Data Ethics: They support implementing data ethics in digital government projects, emphasizing the placement of trust at the core of design and delivery and upholding public integrity through specific actions by governments and public organizations.
- OECD Recommendation on Enhancing Access to and Sharing of Data: Internationally agreed set of principles and policy guidance on how governments can maximize the cross-sectoral benefits of many types of data, while protecting stakeholders' rights.
- UN OHCHR Human rights-based approach to data: A set of principles, recommendations and good practices to guide data collection and disaggregation from a human rights lens.
- UN System Chief Executives Board for Coordination (CEB): Principles for the Ethical use of AI by the UN System.

Box 3. FOCUS: UN System Chief Executives Board for Coordination (CEB) Data Governance Principles

1. Value

- Maximizing the Value of Data: Emphasize fostering a culture that values data as a crucial asset to foster development for all. This includes promoting data quality, responsible data use, and enhancing interoperability through standardized definitions and classifications.
- Enabling Data Use and Reuse: Encourage the appropriate access, sharing, and reuse of data across borders and sectors, ensuring that data contributes to public good. Embed principles of mutuality and solidarity into data governance to ensure that the value of data benefits both individuals and society as a whole, rather than being exploited for private profit alone.
- Interoperability: Adopt standardized formats, standard metadata, and data definitions to ensure seamless data exchange and collaboration between systems and across borders.
- Building a Data-Literate Society: Promote data literacy and access to technology to empower individuals and organizations to effectively use, analyze and understand data. Set educational initiatives and infrastructure that enable people to work with and comprehend data responsibly.

2. Trust

- Human Rights-Based Approach to Data: Ground all data governance practices in international human rights frameworks, ensuring privacy, protection, and the security of personal data, particularly for vulnerable groups. Prioritize safeguarding access to personal data before enabling its use and reuse, ensuring that data governance respects fundamental rights before facilitating innovation.
- Accountability and Transparency: Establish clear roles, responsibilities, and oversight mechanisms to ensure accountability in data governance. This includes transparent decision-making processes and the provision of redress mechanisms for individuals and communities affected by data misuse. Also, the enforcement of due process limitations to regulate the lawful access, processing and use of data.
- Data Quality and Security: Implement robust measures to ensure data quality and protect the security of data and its supporting infrastructure throughout the data lifecycle. This includes the use of context-aware data quality assessments and safeguarding against data corruption or breaches. It also entails the prevention of unauthorized re-identification and the protection of individuals and social groups from discrimination resulting from data processing and analytics.

3. Equity

- Promoting Equity in Data Governance: Ensure that the benefits of data are equitably distributed, focusing on reducing data poverty and preventing discrimination. This principle emphasizes participatory decision-making processes, ensuring that marginalized individuals and communities have control over their data and are actively involved in shaping data policies and governance structures. It also promotes representation in governance bodies and community-driven data initiatives to ensure that data governance reflects the needs of diverse populations.
- Digital Self-Determination: Advocate for individuals' rights to control their personal data, enabling them to make informed decisions about its use and ensuring their agency in the data ecosystem. It promotes clear, accessible consent mechanisms and opt-in frameworks that allow users to meaningfully engage with how their data is used.
- Fairness and Non-Discrimination: Promote fair treatment in data collection, analysis, and use, actively working to mitigate bias and prevent discrimination in data practices. Ensure algorithmic transparency and explainability, requiring that decision-making models be interpretable, accountable, and auditable to prevent opaque systems that disproportionately impact specific communities.
- Data Stewardship and Ethical Reuse for Public Good: Establish strong data stewardship practices that balance data protection with responsible, ethical reuse to maximize societal benefits. Ensure data is managed responsibly, ethically, and securely to support public interests while upholding individuals' rights. Encourage open data initiatives and collaborative governance models that provide fair access to data for research, policy-making, and social innovation while respecting privacy and security concerns.

Assessment Questions & Rationale:

Principles (Documentation)

1. **Question: Are data-related principles (e.g., transparency, accountability, fairness, participation) officially documented or publicly endorsed at the political or organizational level?**
   - **Rationale**: This question assesses whether the organization has established high-level, often ethically or politically significant, principles guiding its overall approach to data. Formal documentation and leadership endorsement signal commitment and provide a basis for trust and public accountability. A positive answer here represents the first step of defining core values for data use.
2. **Question: Are separate technical or operational data governance principles (e.g., data quality, stewardship, access control) formally defined or adopted at the agency/department level?**
   - **Rationale**: While high-level principles set the vision, operational principles translate that vision into practical guidance for day-to-day data management. This question checks for the existence of specific rules or standards related to how data is handled by practitioners. Formal definition ensures consistency and clarity for data creators, users, and stewards. A positive answer indicates these necessary practical guidelines have been established.

Principles (Implementation & Alignment)

3. **Question: Are these principles embedded into operational practices across data lifecycle stages (e.g., collection, processing, sharing, reuse)?**
   - **Rationale**: Principles are only effective if they are put into practice. This question evaluates the critical link between documented principles and actual implementation. Embedding means principles influence workflows, tools, training, and decision-making throughout the data's lifecycle and processing. A positive answer here indicates that the principles are embodied in actual data processing practice.

4. **Question: Are the adopted principles aligned with international or national frameworks and best practices (e.g., FAIR, CARE, African Union, OECD principles)?**
   - **Rationale**: Alignment with recognized external standards and regulations is crucial for interoperability, compliance, ethical responsibility (especially for specific data types like Indigenous data), and demonstrating good practice internationally or nationally. This question assesses whether the internal principles are consistent with broader legal, ethical, and technical data ecosystems. A positive answer shows consideration for external best practices and requirements.

5. **Question: Are these principles enforceable through contracts, MOUs, or data-sharing agreements with external partners?**
   - **Rationale**: Data governance extends beyond organizational boundaries when data is shared or received. This question assesses whether the organization ensures its data principles (particularly those related to security, privacy, use limitations, and quality) are contractually binding on external parties. Enforceability is vital for managing risk, maintaining control over shared data, and ensuring partners adhere to necessary standards. A positive answer indicates that principles are legally integrated into external data collaborations.

Methods

Defining the principles that underpin a data governance framework is not just a technical exercise—it establishes core values. Principles determine how decisions are made, data is handled, and trust is built. They shape what is considered fair, legitimate, and effective in data governance processes. While defining clear purposes sets the direction, articulating principles ensures that the journey toward those purpose reflects the values and priorities of those affected.

To define these principles thoughtfully and to be inclusionary, a range of methods can be employed, each offering unique insights into which values should guide governance.

Grounding Principles in Real-World Needs and Perspectives

- The first step is understanding who the governance framework will affect and whose voices should be reflected. **Stakeholder mapping and engagement** are essential here. By identifying individuals and groups with a stake in how data is used—from government agencies and private firms to community organizations and marginalized populations, their concerns and aspirations can begin to be identified. Engagement activities, such as interviews, workshops, or citizen panels, help clarify what principles matter most: Is transparency a key demand? Is inclusiveness being neglected? Are there calls for more accountability or stronger protections against harm? Example: HUD Exchange's Community Engagement Toolkit
- Public-facing initiatives like data assemblies or citizen juries can further ensure that data principles are not just imposed from the top down but co-developed with those most impacted. These participatory processes democratize the definition of principles and anchor them in lived experience. Example: The Data Assembly

Surfacing Values Through Deliberation and Design

- Another tool involves the use of **Data Ethics Canvases**—structured tools that prompt teams to reflect on the ethical dimensions of data use. Based on human-centered design, these canvases guide discussions around risk, power imbalances, and stakeholder impact, helping to crystallize principles such as transparency, non-discrimination, and user agency in concrete terms. They are particularly useful during the early stages of governance design, when teams need to make core values explicit. Example: Data Ethics Canvas
- **Norm engineering workshops** provide another valuable approach. These sessions bring together diverse stakeholders—technical experts, legal advisors, civil society representatives—to debate and reconcile tensions and trade-offs (such as between openness and privacy or innovation and precaution). The goal is to co-construct principles that are not only meaningful but also realistic and adapted to institutional and societal constraints. Example: Norm Engineering

Anchoring Principles in Rights and Responsibilities

- To ensure that data governance principles reflect internationally recognized norms, some organizations conduct **Human Rights Impact Assessments** (HRIAs). This method examines how a data initiative might affect fundamental freedoms—such as the right to privacy, equality, or access to information—and distills key principles that should be embedded to prevent harm and uphold justice. A good practice in assessing legal frameworks through an HRIA is to apply the **three-part test**—which ensures that any interference with rights is prescribed by law, compatible with human rights standards, and necessary and proportional. This test helps guarantee that legal provisions align with human rights principles and maintain a balanced approach. Example: Human Rights Impact Assessment

Engaging the Public and Planning for the Future

- Data governance principles should not only serve institutional needs, they should also **resonate with society**. To this end, tools like **social license assessments** and **civic chartering** offer meaningful ways to gauge public acceptability and co-create normative frameworks. Through citizen juries, focus groups, or co-drafting exercises, communities can articulate which values they believe should govern data use—often surfacing priorities such as dignity, consent, and fairness in new and localized ways. Example: Social License Toolbox and Civic Chartering and Toolkit on Digital Transformation for People-Oriented Cities and Communities

From Principles to Practice

- Even the best-defined principles can fall short if not operationalized. As noted earlier, tools such as **Balanced Scorecards and principles-to-practice mappings** help embed principles into institutional routines and accountability structures. They ensure that values like inclusiveness or proportionality translate into measurable outcomes and concrete procedures—from data sharing agreements to audit trails and metadata standards. See Internal Audit Matrix

Resources for further reading:

- ITU- World Bank: Digital Regulation Platform's Navigating Data Governance: A Guiding Tool for Regulators This article provides practical guidance to ICT regulators, other regulatory agencies (including data protection authorities), and stakeholders dealing with data governance, in monitoring and guiding organizations' data governance practices, focusing on data classification, data interoperability, data availability, quality and

integrity, data access and sharing, and data security and data protection and privacy

- United Nations Chief Executives Board's Proposed Normative Foundations for International Data Governance: Goals and Principles
- United Nations Statistics Division's Fundamental Principles of Official Statistics, Implementation Guidelines
- One Trust Data Guidance
- The State of Open Data Policy Repository
- Transparency and Accountability as Trust Builders in the African Data Governance Ecosystem
- The Danish Institute for Human Rights: Guidance on Human Rights Impact Assessment of Digital Activities
- African Union Convention on Cyber Security and Personal Data Protection

III. WHO

People and Processes in Data Governance

NicoElNino/Shutterstock.com

Core Objective	Why Principles Matter
To identify the key stakeholders and institutional roles responsible for data governance, and to ensure that effective coordination and accountability mechanisms are in place.	Data governance requires a clear delineation of roles and responsibilities, ensuring that appropriate people and processes are in place to make decisions, oversee compliance, and manage data throughout its lifecycle. Effective governance frameworks should establish clear responsibilities, and can also create mechanisms for coordination, compliance, and ethical review to ensure transparency and accountability across all processes.

Taxonomy of Processes, Roles, and Responsibilities

To ensure efficient and transparent governance, the following key roles and processes are often part of a robust data governance framework (non-exhaustive list):

1. Development, Coordination, and Oversight:

- These roles are responsible for setting strategic direction, ensuring alignment with organizational goals, and overseeing the implementation of data governance policies.
- Examples:
  - Board of Trustees or Coordinating Committee—provides overall governance and strategy direction.
  - Chief Data Officer (CDO)—leads the development and execution of the data governance strategy and ensures alignment with broader organizational objectives. The CDO may chair national data governance board or committee, and there are more often Chief Data Officers or coordinators in relevant ministries and other government institutions.
  - Independent Data Regulators—create the necessary framework for data governance that applies to government agencies and/ or private sector organizations.
  - Inter-Ministerial Task Teams—Within government, there may be ministries or agencies where data issues regularly form a key part of their work. For example, Ministries of Trade handle trade in data; Ministries of Justice handle judicial, crime and imprisonment data; Ministries of Investments, Labor & Industry handle economic data. Accusations relating to data theft with security implications may be handled by national security agencies. Coordination among different intergovernmental departments and agencies is essential to provide a coherent data governance framework.

2. Compliance and Adjudication:

- These roles ensure that the organization complies with data protection laws and resolves issues related to data misuse or breaches.
- Examples:
  - Data Protection Officer (DPO)—ensures compliance with data

privacy regulations and manages any regulatory reporting requirements (similar functions may also be performed by lawyers e.g. Non-Disclosure Agreements).

  - Ethics Officers may also have some responsibilities with regards to unethical disclosures of data, whistleblower protection etc.

3. Facilitation and Management:

- These roles are responsible for day-to-day data management, ensuring that data is properly handled, stored, shared, and re-used in a systematic, ethical, and responsible manner.
- Examples:
  - Data Stewards—manage and maintain the quality, security, and responsible access and re-use of data across different domains within the organization and with third parties.
  - Data Custodians—responsible for capturing, storing and disposing of data for technical tools for data-provisioning. They work with the Data Stewards to ensure data quality.
  - Archivists—manage the long-term preservation of data and documents, ensuring that historical records and critical datasets are securely stored, accessible, and properly maintained for future use.
  - Maturity Assessors—evaluate the organization's data management capabilities against established data governance maturity models.

4. Review and Guidance:

- These roles provide ethical oversight and guidance to ensure that data governance processes uphold moral and ethical standards and complies with legal frameworks.
- Examples:
  - Ethical Review Board— evaluates and guides decisions involving sensitive data use, ensuring adherence to ethical principles and frameworks.
  - Compliance Team— reviews compliance with existing legal and organizational frameworks.

Establishing Data Governance Decision Provenance

Documenting decision-making authority is critical for transparency and accountability. By identifying who is involved in each stage of the data lifecycle, organizations can ensure that decisions are traceable and aligned with governance standards.

Key roles to define include:

- Responsible: Who is responsible for making the decision? Who is subordinate? Who is responsible for implementing the decision?
- Accountable: Who is ultimately accountable for the decision's outcome?
- Consulted: Who needs to be consulted before a decision is made?
- Informed: Who should be informed about the decision after it is made?
- Defining the roles involved through stakeholder mapping ensures clarity and accountability in governance decisions across the entire data lifecycle—from data collection and processing to share and use.

Table 4. Stakeholder mapping

RACI	Planning	Collection	Processing	Sharing	Analyzing	Using
Responsible						
Accountable						
Consulted						
Informed						

Assessment and Questions

Leadership and Strategic Mandate

1. Question: Is there a designated Chief Data Officer (CDO), Chief Data Protection Officer (DPO), Chief Data Steward, or equivalent senior data leader with a formal mandate and adequate resources?

- **Rationale**: Having a senior data leader signals strategic commitment and ensures accountability for data governance across the organization. This role provides the authority and resources needed to implement data initiatives, coordinate across departments, and represent data priorities at the leadership level.

Role Clarity and Accountability

2. Question: Are the responsibilities of keydata governance roles (e.g., CDO, DPO, AI ethics board, data steward) clearly defined and documented?

- **Rationale**: Clarity around roles reduces confusion, prevents gaps or overlaps in responsibility, and supports more efficient implementation of policies and safeguards. Clearly defined responsibilities also facilitate onboarding, training, and cross-team collaboration.

3. Question: Are formal or informal roles or structures in place for data stewardship, privacy oversight, and ethical data use within public se ctor institutions (e.g., data stewards, privacy officers, ethics boards)?

- **Rationale**: The presence of these roles—whether formally institutionalized or informally recognized—indicates an organization's readiness to manage ethical, privacy, and stewardship responsibilities. These structures help ensure that data use is not only legal but also responsible, ethical, and aligned with societal values.

13. **Question: Is there an independent data protection regulator or authority in place to oversee data privacy and protection practices, ensuring compliance with national and international data protection laws and upholding human rights?**

- **Rationale**: An independent regulatory authority ensures that data privacy and protection practices are free from political, economic, or other undue pressures, maintaining the integrity of data governance. These regulators play a critical role in holding organizations accountable to established laws, conducting periodic reviews, and ensuring that governance structures comply with international human rights standards.

Inter-Institutional Coordination

15. **Question: Is there any collaboration (structured or unstructured) on data governance between government agencies (e.g., inter-agency data committees, joint strategies, shared data platforms)?**

- **Rationale**: Data often spans multiple departments. Inter-developmental or inter-agency collaboration enables a coherent and unified approach to governance, improves data interoperability, and reduces redundancy. It's also essential for addressing shared challenges like crisis response or digital transformation.

Multi-Stakeholder Engagement

17. **Question: For governments, is there any coordination (formal or informal) on data governance with actors outside government (e.g., academia, civil society, private sector)?**

- **Rationale**: Engaging external stakeholders enhances the legitimacy, quality, and impact of data governance frameworks. It brings in diverse expertise, helps address blind spots, and ensures that governance reflects public needs and values—especially important for building trust and enabling data for public good.

Methods and Tools

Establishing clear roles and responsibilities is essential for effective and accountable data governance. Without clarity on who is responsible for what, efforts can become fragmented, leading to inefficiencies, duplicated efforts, or gaps in oversight. The following examples are some existing methods and tools that can help identify key actors, assign responsibilities, and foster coordination across the data ecosystem:

- Decision Provenance RACI Model: This tool helps define and document decision-making roles by identifying who is Responsible, Accountable, Consulted, and Informed at each step of the data governance process.
- Map your data ecosystem: This tool helps policymakers identify and understand the key actors, data sources, and interactions within a data ecosystem. By mapping data flows, infrastructure, and value exchanges across public institutions, private organizations, and communities, it reveals opportunities to improve data management, accessibility, and governance.
- Mapping of Existing Data Governance Actors: A comprehensive mapping exercise to identify all current actors involved in data governance, their roles, and their responsibilities within the organization. This ensures clarity on who is involved in various stages of data management and oversight.

Resources for further reading:

- ITU Navigating Data Governance: A Guiding Tool for Regulators.
- ITU Technical Report D4.1 – Framework for Security, Privacy, Risk, and Governance in Data Processing and Management

IV. WHAT

The policies, practices, and technologies that govern each stage of the data lifecycle, ensuring that data is handled with purpose and aligned with guiding principles.

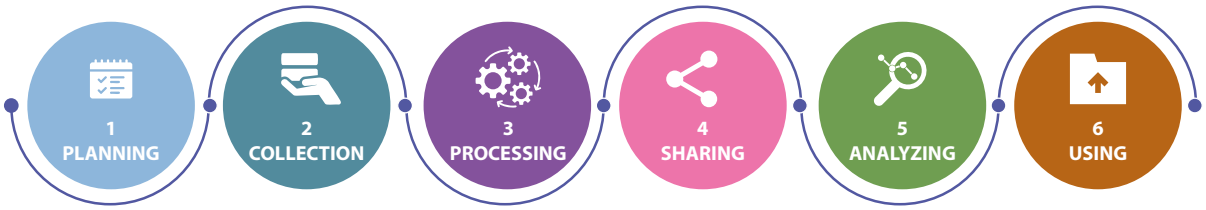
Ar_TH/Shutterstock.com

Core Objective

To ensure that data is governed effectively across its entire lifecycle— from planning and collection to use and reuse—so that it fulfills its intended purpose(s) while aligning with established principles, ethical standards, and regulatory compliance.

Analyzing the Data Lifecycle?

As previously mentioned, the data lifecycle refers to the various stages data goes through—from its initial planning to its ultimate use in decision-making. While different frameworks may highlight different stages or use varying terminology, the most commonly recognized phases include:



- **Planning** – Identifying data needs, intended uses, and governance requirements.
- **Collection** – Gathering data through surveys, sensors, transactions, mapping, imagery or other means.
- **Processing** –Cleaning, validating, organizing, storing, and preserving data for use, including deletion when necessary, and ensuring proper handling throughout the data lifecycle.
- **Sharing** – Making data accessible to others, whether through platforms, APIs, or data collaboratives.
- **Analyzing** – Interpreting the data to generate insights.
- **Using** – Applying those insights to inform decisions, policies, or services.

At each stage, several concerns and considerations must be addressed to ensure data governance aligns with principles and achieves the specified purposes. The section below highlights these considerations, along with recent developments and tools.

Box 4. 10 Data Governance Mechanisms

To implement data governance principles and decisions across the data lifecycle, multiple mechanisms can be considered. It does not include legislation or regulation such as data protection or privacy law as these serve as overarching legal frameworks rather than operational or procedural tools for day-to-day governance. Among them:

1. **Contractual mechanisms:**
   - Legally binding agreements that set terms and assign responsibilities for data access, sharing, usage, and limitations on third-party interaction with data access and use, rights of third parties, control of technical mechanisms to access data such as APIs, etc. They can be tailored to the specific case or set in general terms
   - Examples: **Data Sharing Agreements (DSAs), Memorandums of Understanding (MOUs), Service-Level Agreements (SLAs), End-User License Agreements, API terms of service, API Usage Policies**

2. Policies & Guidelines

- Institutional and governmental guidelines that outline how principles for data governance should be implemented.
- Examples: **Open Data Policies, AI Ethics Guidelines**.

3. Technology & Governance by Design

- Technical solutions embedded within systems to enforce governance principles.
- Examples: **Differential privacy, federated learning, encryption, access controls, technical protection measures (TPMs)**, **data architectures and integration design.**

4. Standards and Vocabulary

- Common protocols and definitions to document and ensure data quality, security, interoperability, consistency, and usability.
- Examples: **F ISO 27001 (Information Security), DCAT (Data Catalog Vocabulary)**.

5. Codes of Conduct

- Voluntary or mandatory frameworks guiding responsible data use.
- Examples: **EU Code of Conduct on Disinformation**.

6. Procurement & Vendor Management

- Integrating data governance requirements into procurement processes.
- Examples: **Public sector data-sharing requirements in vendor contracts**.

7. Licensing

- Mechanisms defining permissions for data reuse and distribution.
- Examples: **Creative Commons Licenses, Open Data Licenses**.

8. Data Stewardship & Institutional Arrangements

- Establishing roles and responsibilities for managing data in alignment with governance goals
- Examples: **Chief Data Stewards, Data Trusts, Independent Auditors**.

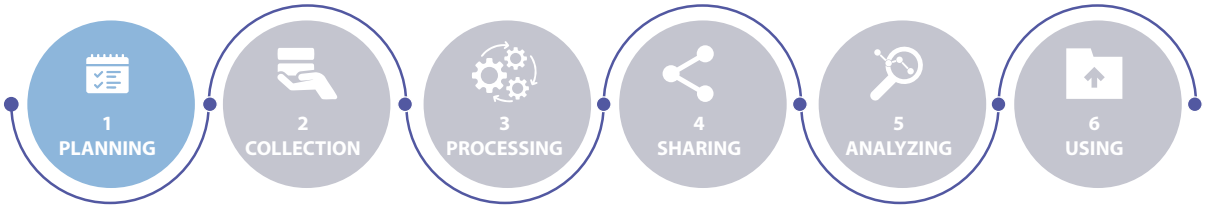
9. Audit & Compliance Mechanisms

- Methods to monitor and enforce adherence to governance policies.
- Examples: **Impact Assessments, Compliance Audits, Algorithmic Transparency Reports.**

10. Training and Cultural Change Initiatives

- Mechanisms to embedding data training and cultural change in the organization, such as mandatory training courses (privacy, data security) in-person or virtual sessions, and data governance training.

1. Planning



- In the planning stage, decisions are made that define the scope, purpose, feasibility, and governance of the data initiative. These early steps fundamentally shape the success and integrity of data use later on. Effective planning requires understanding the intended public value of the data, mapping key actors and systems, identifying governance gaps, and designing processes that are aligned with legal, ethical, and contextual realities.
- This stage also involves conducting early risk and cost-benefit assessments, defining roles and responsibilities, and establishing governance models that will be implemented across the data lifecycle. Particular attention should be paid to aligning with the principles of purpose limitation, data minimization, interoperability, and trustworthiness—especially in environments involving cross-sector or cross-border collaboration.

Table 5. Checklist: Planning

Task	
Purpose and value of data clearly defined and documented	☐
Stakeholders and affected communities identified and mapped	☐
Legal and policy landscape reviewed, including prior efforts	☐
Governance model designed (roles, responsibilities, decision-making)	☐
Scope, goals, and limitations of the data project articulated	☐
Financial, technical, and human resources evaluated and secured	☐
Data sharing agreements (MoUs, DSAs) and legal templates prepared	☐
Interoperability needs and shared vocabularies discussed	☐
Risk assessments conducted (e.g., privacy, security, AI risks)	☐
Stakeholder engagement and trust-building strategy in place	☐
Communication and transparency plan drafted	☐
Feedback loops and evaluation metrics defined	☐
Data preservation and storage needs assessed and planned	☐

10 Assessment Questions

1. Is there a clearly defined purpose(s) for collecting and using the data?

Rationale: Establishing that the purpose is aligned with societal goals and principles such as purpose limitation, helping prevent function creep or misuse of data.

2. Have you identified and planned for any ethical concerns or potential public backlash?

Rationale: Anticipating societal and ethical issues helps secure a social license to operate and maintain public trust.

3. Are all relevant stakeholders, including marginalized groups, identified and engaged?

Rationale: Inclusion improves data quality, enhances legitimacy, and ensures that governance frameworks reflect diverse values and needs.

4. Have assumptions, success metrics, and decision-making criteria been documented?

Rationale: Clarity on assumptions and evaluation metrics supports accountability and helps adapt the project in response to new information.

5. Are there documented legal agreements covering data access, sharing, and privacy?

Rationale: Legal clarity builds trust among partners and ensures compliance with data protection regulations.

6. Have you evaluated the costs and benefits of this data initiative?

Rationale: Sound financial and risk planning ensures that data investments deliver expected public value while mitigating potential harms.

7. Are there technical requirements identified for interoperability and scalability?

Rationale: Planning for data integration and scalability avoids bottlenecks and supports long-term use of the data across sectors.

8. Has a shared vocabulary or taxonomy been agreed upon by stakeholders?

Rationale: A shared language ensures common understanding and enables data discoverability and reusability.

9. Are there procedures for regular review and adaptation of the data governance strategy?

Rationale: Governance needs to be iterative and responsive to emerging risks, opportunities, and stakeholder feedback.

10. Have past efforts or lessons learned from benchmarking similar data initiatives been reviewed?

Rationale: Reviewing past initiatives helps avoid duplication, incorporates best practices, and builds on existing infrastructure or trust relationships.

Methods and Tools

- Open Data Impact Framework: Periodic Table of Open Data Elements detailing the enabling conditions and disabling factors that often determine the impact of open data initiatives.
- Data Management Frameworks: Use frameworks like DAMA-DMBOK (Data Management Body of Knowledge) to guide standardized, transparent management practices from the start.
- Workflow Tools: Design process maps and decision flows to assign responsibilities, structure decision processes, and clarify the intended data, such as the 4Ps model.
- Legal Templates: Develop early drafts of data sharing agreements (DSAs), Memoranda of Understanding (MoUs), and Privacy Impact Assessments (PIAs) to formalize partnerships and ensure compliance with data protection laws.
- Privacy and Security Risk Assessments: Conduct initial assessments using tools like Data Privacy Impact Assessments (DPIAs) to identify and mitigate privacy risks.

- Shared Glossaries and Vocabularies: Develop a shared vocabulary early on with glossaries, taxonomies and knowledge management systems to ensure that all partners are aligned on key terms and concepts.
- Upcoming UNESCO Public Data Classification (forthcoming).

Resources

UNDP Model Governance Framework for Digital Legal Identity System

ITU Navigating Data Governance: A Guiding Tool for Regulators

ITU Technical Report D4.1 – Framework for Security, Privacy, Risk, and Governance in Data Processing and Management

Singapore Infocomm and Media Development Authority (IMDA) Guide to Data Valuation for Data Sharing

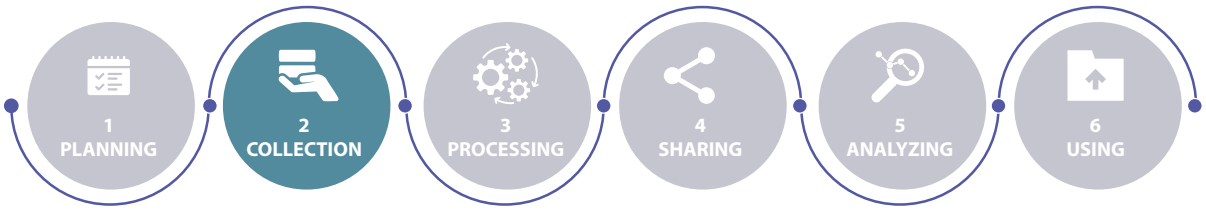
Box 5. Different types of data

Different types of data require distinct governance considerations to ensure proper handling and compliance with legal and ethical standards.

- **Personally identifiable information (PII)**, such as names, addresses, and social security numbers, demands stricter governance due to privacy and security risks. Safeguarding PII involves implementing stringent access controls, encryption, and adhering to data protection instruments such as the Council of Europe's Convention 108, which has been ratified by many countries outside the European space. Differentiating governance strategies based on the sensitivity of the data helps to protect individual rights while maximizing the utility of less sensitive information for innovation and public good. Anonymization of data (e.g. through masking, data aggregation or pseudonymization) can help deal with some of the difficulties; however, anonymization needs to be handled carefully, as it has been pointed out that identity can be inferred (e.g. through linkage attacks, cross-referencing or linking data, especially correlated data. Pseudonymization is often assumed to mean anonymization of data, but this is not always effective - Why Are Legacy Data Anonymization Techniques Failing?
- In contrast, **non-personal data**, such as aggregated or anonymized datasets, or datasets pertaining to objects or locations (e.g., data about buildings, supplies of medicine, etc.) while posing fewer privacy concerns, still requires careful governance to maintain data quality, ensure ethical use, and foster transparency.
- **Unstructured data**, as it lacks a predefined format or organizational framework, requires specialized tools and strategies to unlock its potential. From a planning perspective, this entails different challenges such as data discovery and classification, data integration, scalability and storage considerations, privacy and security compliance as well as resource allocation to manage all the potential challenges.

It is however important to note that the level of governance should be based on the sensitivity of the data and not just on data types. For instance, non-personal data can also be highly sensitive due to their business value, ethical implications, or regulatory requirements.

2. Collection



- The collection stage determines which what data will be gathered, from whom, by which means, and under which legal and ethical conditions. It is the frontline for ensuring responsible and lawful data use. Good data governance at this stage focuses on legality (compliance with national and regional laws and regulations), proportionality (collecting only what is needed for specific purposes), consent (individual and/or collective), transparency (informing those affected), and equity (inclusive representation in data sources), and ethics (ensuring data use aligns with moral principles such as fairness, accountability, and respect for rights).
- Traditionally in terms of statistics, there was a trade-off between the amount of data collected and the accuracy of statistical results (e.g. for inferring population characteristics from a limited sample size), as well as the costs of data collection (e.g. cheaper sampling methods versus population census). Considerable priority in statistics (as well as Statistics Offices and consultancies) has been given to reducing sampling costs, while seeking to preserve representativeness and accuracy. However, recent techniques, inferred data and machine analysis can help address some of these trade-offs.
- Data should be collected in a manner that reflects the purpose and respects data protection obligations. At the same time, practical considerations such as interoperability, metadata documentation, the implementation of clear norms and standardized methodologies for data collection and bias identification are essential to ensure the usability and trustworthiness of the data later on.

Table 6. Checklist: Collection

Task	
Data collection aligns with stated purpose and minimization principles	☐
Consent mechanisms established and documented (including dynamic consent)	☐
Marginalized or underserved populations are represented	☐
Privacy-by-design principles applied in collection systems	☐
Data formats standardized and interoperable	☐
Metadata and documentation practices defined	☐
Cross-border data flow and localization requirements assessed	☐

Tools in place to detect and mitigate bias during collection	☐
Encryption or anonymization methods applied at point of collection	☐
Legal obligations reviewed and adhered to (e.g., lawful basis, jurisdictional issues)	☐
Feedback mechanisms in place to adapt collection processes if issues arise	☐

10 Assessment Questions with Rationale

1. Do the data collection methods and practices follow the principle of proportionality (only what is necessary)?

Rationale: Minimizing data collection may reduce privacy risk, lower processing costs, and align with legal requirements

2. Has meaningful and informed consent been obtained and recorded where required?

Rationale: Consent builds legitimacy and meets legal obligations—especially important for sensitive or personal data.

3. Are the rights and perspectives of groups in situation of marginalization considered in data collection design?

Representation ensures the data supports policymaking for all and avoids reinforcing structural biases.

4. Is there a privacy-by-design framework applied to data collection tools or platforms?

Rationale: Embedding privacy from the outset strengthens protections and prevents costly retroactive adjustments.

5. Are metadata and contextual information being collected alongside raw data?

Rationale: Metadata is critical for evaluating relevance, quality, and enabling responsible future reuse.

6. Are standardized formats and protocols used to support data interoperability?

Rationale: Interoperability enhances the potential for collaboration, cross-sectoral reuse, and system integration.

7. Are there safeguards (encryption, anonymization) applied at the point of collection?

Rationale: Early-stage protection of sensitive data reduces exposure to breaches and supports secure lifecycle handling.

8. Are tools or methods in place to detect potential bias during data collection?

Rationale: Early identification of bias allows course correction and enhances the fairness of later analysis.

9. Are local and international data protection rules considered in cross-border collection?

Rationale: Understanding jurisdictional requirements helps ensure compliance and avoid legal risk when data crosses borders.

10. Is the data collection process regularly reviewed and updated to ensure compliance with legal obligations?

Rationale: Continuous review allows for correction of oversights and adapts practices to new risks or opportunities.

Methods and Tools

Practices for Collection

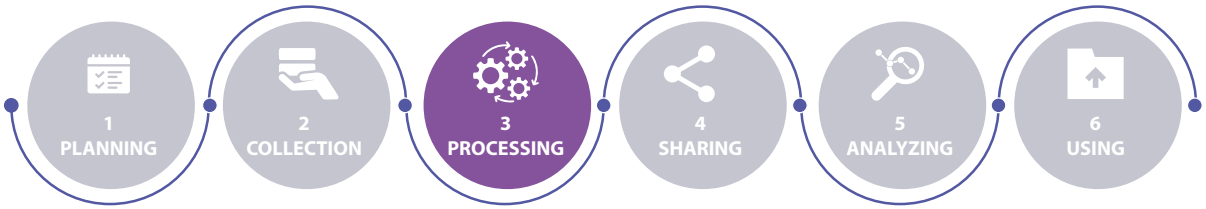
- Data Minimization Tools: Implement systems that ensure only essential data is collected, using techniques such as differential privacy to minimize unnecessary data collection.
- Consent Management Platforms: Use tools to obtain, manage, and record meaningful consent from data subjects, ensuring that consent processes are transparent and well-documented.

- Privacy Impact Assessments (PIA): Regularly conduct PIAs to evaluate potential privacy risks and ensure that data collection processes align with ethical and legal standards. For instance, use the RD4C decision-making tools to identify and address risks in data collection. These frameworks can guide privacy-conscious design even outside the child data context.
- Collective Intelligence Data Collection: Especially for public sector organizations, the implementation of participatory collection methods such as crowdsourcing, citizen science, and sensor-based contributions can enhance data completeness and legitimacy, especially for environmental, health, and community-level datasets. However, the use of online survey methods and sampling can also introduce some bias (e.g. it can be difficult to make interferences about elderly, sick or offline populations from online surveys).
- Privacy-by-Design Frameworks: Adopt frameworks that integrate privacy considerations into the design of data collection systems from the outset, ensuring robust privacy protections throughout the lifecycle.
- Regulatory Compliance Tools: Use compliance systems to ensure adherence to relevant local, national, and international data protection laws.
- Evaluating Data Assets
- Data Curation Tools: These can help to define, meet and support the adoption of appropriate standards.
- Data Inventory Tools: Use tools and platforms to audit and catalog existing datasets, ensuring that all available data assets are appropriately tracked and organized.
- Data Context Analysis: Implement tools to assess the relevance, accuracy, and applicability of data in relation to the specific use case or problem being addressed.
- Data Quality Management Tools: Use systems to evaluate and maintain data quality, ensuring completeness, consistency, and accuracy throughout the data lifecycle.
- Bias Detection Tools: Apply bias detection algorithms and consult external experts to identify and address potential biases in collected data.
- Data Safeguards and Documentation: Implement platforms that document the data collection process, including metadata, limitations, and biases, to promote transparency and accountability.

Resources

- Royal Society: From privacy to partnership: the role of Privacy Enhancing Technologies in data governance and collaborative analysis
- ITU Navigating Data Governance: A Guiding Tool for Regulators
- UNESCO Data-invisible groups and data minimization in the deployment of AI solutions: policy brief

3. Processing



The processing stage involves preparing collected data for storage, analysis, or exchange. This includes tasks like cleaning, categorizing, transforming, securing, and validating data. At this stage, governance decisions can help determine how well data retains its integrity and utility, how securely it is managed, and how well it complies with legal and ethical standards.

Data processing often takes longer and proves more arduous than initially planned, despite the greater availability of computing power and resources. It involves high-level considerations of who and why, as well as more mundane, detailed considerations (e.g. offsite/onsite processing capacity, bandwidth, transmission capabilities, power, back-up, records management, storage).

Data processing should strike a balance between cost, usability and protection, ensuring the data is fit for purpose while minimizing risks such as re-identification, unauthorized access, or misuse. It also includes decisions on system architecture (e.g., centralized vs. decentralized storage), data interoperability, and long-term scalability.

Policies for the preservation of public records and data, as well as structured guidelines for archiving and retaining data, should be established to ensure long-term accessibility and compliance with legal and regulatory frameworks.

Table 7. Checklist: Processing

Task	
Data cleaning and transformation processes documented	☐
Data quality (accuracy, consistency, completeness) validated	☐
Data categorized and classified (e.g., sensitive vs. non-sensitive)	☐
Data provenance and version history maintained	☐
Encryption and access controls implemented	☐
Internal access protocols defined (e.g., tiered access, audit logs)	☐
Privacy-enhancing technologies considered or applied	☐
Backup, archival, and effective and secure deletion systems in place	☐
Data preservation policies established (including public records)	☐
Archival systems and long-term storage mechanisms in place	☐

Compatibility with interoperability standards ensured	☐
Processing activities reviewed for legal and ethical compliance	☐
Mechanisms in place to avoid unintended data linkages	☐

10 Assessment Questions

1. Has the data been cleaned, validated, and made analysis-ready?

Rationale: Clean and validated data ensures reliability of downstream insights and avoids propagating errors.

2. Is the data categorized by sensitivity and use-case (e.g., PII, public, restricted)?

Rationale: Categorization enables appropriate protection measures and responsible reuse.

3. Is data provenance (origins, transformations, and versions) tracked?

Rationale: Provenance promotes transparency, reproducibility, and accountability, especially for AI systems.

4. Are robust security protocols in place (e.g., encryption, anonymization, access control)?

Rationale: Security ensures that sensitive or valuable data is protected from breaches or unauthorized modification.

5. Are internal access controls role-based and auditable?

Rationale: Tiered access and audit logs enhance security and enable accountability by tracking data usage.

6. Are privacy-enhancing technologies (PETs) considered or implemented during processing?

Rationale: PETs such as homomorphic encryption or federated learning reduce privacy risks while allowing data use.

7. Is the data processing environment protected against unauthorized access or system failure?

Rationale: Secure processing environments reduce exposure to breaches and ensure business continuity.

8. Are back-up, preservation, archiving, and secure deletion protocols clearly defined and followed?

Rationale: Clear protocols for preservation, archiving, and secure deletion are vital for legal compliance, data integrity, and risk reduction. Proper retention ensures long-term access to critical data, while secure deletion minimizes exposure and meets regulatory requirements.

9. Are processing formats and methods compatible with anticipated interoperability needs?

Rationale: Using common data formats and standards ensures that processed data can be reused across platforms.

10. Have steps been taken to prevent harmful aggregation or unintended inferences?

Rationale: Even non-sensitive data can become risky when combined—anticipating these risks helps avoid misuse.

Tools and Practices

- Developing Storage and Operations Guidelines
  - Backup and Archival Systems: Implement cloud-based or on-premise backup services that ensure automatic data backups and provide reliable archival solutions.
  - Data Deletion Tools: Utilize secure data deletion tools and methods to ensure complete removal of sensitive data, preventing unauthorized recovery.

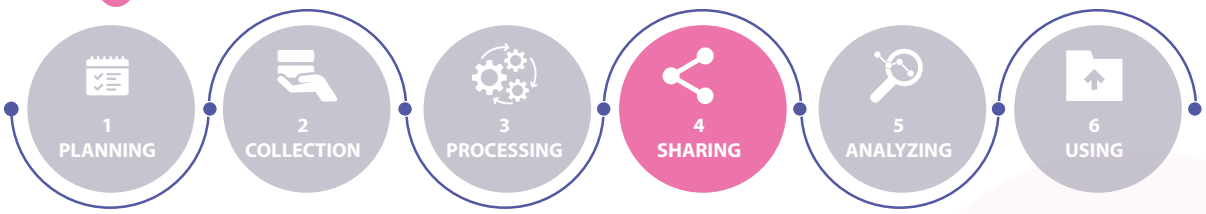
- Device Management: Deploy systems to manage, track, and securely retire devices that store sensitive information.

- Building Robust Security
  - IT Security Tools: Implement advanced security measures such as firewalls, intrusion detection systems, and encryption technologies to safeguard data from unauthorized access and breaches.
  - Security Expertise: Regularly consult with internal and external data security experts to ensure robust protection and compliance with best practices.
  - Security Training: Conduct training programs for staff to enhance awareness and preparedness in handling potential data security incidents.
  - Crisis Communication Strategy: Develop and maintain incident response plans to address and communicate effectively during data breaches or security crises.
- Establishing Internal Access & Security Protocols
  - Tiered Access Management: Set up identity and access management systems that provide different levels of access based on roles and responsibilities within the organization.
  - Multi-Factor Authentication (MFA): Use multi-factor authentication processes to enhance the security of access to sensitive data.
  - Audit Trails: Implement logging and audit systems to document and track who accesses data, when, and what changes or actions are taken over time.
  - Data Encryption: Ensure the encryption of sensitive data during processing and transmission to safeguard privacy and security.
- Categorizing, Classifying & Taxonomizing Data (Data Architecture)
  - Data Taxonomy Tools: Implement systems to categorize and classify data, capturing provenance and ensuring proper organization for analysis.
  - Data Mapping Tools: Use data integration and mapping tools to transform datasets into formats suitable for analysis while maintaining accuracy.
  - Documentation Tools: Establish documentation and version control practices to track assumptions, cleaning processes, and other data-handling activities.
  - Data Integrity Tools: Deploy safeguards to ensure data integrity and compatibility when aggregating or correlating different datasets, avoiding processing errors.

Resources

- ASEAN Guide on Data Anonymization
- UN Guide on Privacy-Enhancing Technologies for Official Statistics
- UNESCO Recommendation on the Preservation of Documentary Heritage
- Upcoming UNESCO Public Data Classification (TBC)

4. Sharing



The sharing stage governs how, with whom, and under what conditions data is exchanged. It is a critical stage for trust, interoperability, and legal compliance—especially in public-private partnerships or cross-border collaborations. Access to data for re-use can unlock collective and public value but also raises reputational, legal, and ethical risks if poorly managed.

Good governance at this stage requires clear legal agreements, technical safeguards, transparent communication, and a shared understanding of purpose among partners. It should also account for the diversity of data-sharing models—from open portals to federated exchanges and data collaboratives—and anticipate unintended consequences like overexposure or misuse of combined datasets.

Data sharing and exchange do not refer to the sale of data for financial or commercial purposes, such as the sale of personal data to marketers, advertisers, or third-party vendors (e.g., selling telephone numbers to publicity or spam callers). However, it is important to distinguish between such commercial sales and situations where data access may be provided for a fee—such as for academic purposes, research, or in cases where the data is used for non-commercial purposes that contribute to public good or innovation. These scenarios may still involve data sharing agreements, but they are typically subject to different considerations and regulations, particularly regarding transparency, consent, and the purpose for which the data is used.

Table 8. Checklist: Sharing

Task	
Purpose and scope of data sharing clearly defined and agreed	☐
Legal agreements (e.g., DSAs, MoUs, licensing) formalized	☐
Internal stakeholders aligned and briefed on governance expectations	☐
External communication strategy developed	☐
Reputational risks and mitigation plans identified	☐
Interoperability protocols established (e.g., common vocabularies, formats)	☐
Technical safeguards in place (e.g., secure APIs, federated access)	☐
Privacy and security risks from data combination anticipated	☐
Governance for third-party intermediaries or platforms clarified	☐
Social license and ethical considerations reviewed	☐
Ongoing coordination and review mechanisms established	☐

10 Assessment Questions with Rationale

1. Is there a clearly defined and documented purpose for data sharing?

Rationale: Purpose clarity prevents misuse, ensures relevance, and supports legal compliance such as purpose limitation

2. Are legal agreements in place to govern rights, responsibilities, and access?

Rationale: Contracts protect parties involved, clarify liabilities, and enforce compliance with law and ethics.

3. Have internal stakeholders (e.g., legal, IT, ethics teams) been consulted and aligned?

Rationale: Cross-functional alignment ensures that operational, legal, and reputational risks are properly managed.

4. Are external communications clear on why data is being shared and with whom?

Rationale: Transparency builds public trust and guards against perception of misuse or hidden agendas.

5. Are reputational risks assessed, including those related to the sharing partner or use case?

Rationale: Risk management protects institutional credibility and anticipates public reaction.

6. Are data formats, vocabularies, and access protocols standardized for interoperability?

Rationale: Technical alignment prevents data fragmentation and enhances reuse across systems or sectors.

7. Are privacy, security, and re-identification risks from dataset combination assessed?

Rationale: Combined datasets may inadvertently reveal sensitive insights—these risks should be preemptively managed.

8. Are there clear governance structures for intermediaries or shared platforms (e.g., data trusts)?

Rationale: Third-party platforms should be governed to ensure they don't circumvent rules or erode accountability.

9. Has the social license been considered beyond individual consent?

Rationale: Community or collective expectations (especially in sensitive domains) should be respected to avoid backlash.

10. Is there a plan for continuous evaluation and course correction in the sharing partnership?

Rationale: Ongoing oversight ensures the partnership remains aligned, effective, and adaptive to change.

Box 6. Data Collaboratives

- Data collaboratives offer structured ways for public-private data exchange while addressing trust, privacy, and competitive concerns. Different models allow for varying levels of control and governance:
- Data Pooling: Organizations combine data into a shared pool for collaborative analysis or innovation, but strong governance is essential to ensure privacy and intellectual property protection
- Data Intermediaries: Neutral third parties manage data exchange, enforcing governance rules and ensuring compliance, reducing risk for data-sharing entities.
- Intelligence Sharing: Organizations share insights or analysis derived from their data without exposing the raw data, enabling collaboration while maintaining control.
- Challenges and Prizes: Encourage solving complex problems with shared data through competitions that offer rewards or recognition.
- Research Partnerships: Between public and private sectors support data sharing for academic or policy research, often governed by agreements ensuring ethical handling and proprietary protection.
- Public Interfaces: Provide open access to specific datasets or tools, enabling a wide range of stakeholders—researchers, developers, and civil society organizations—to analyze data and generate insights.

Box 7. FOCUS: Data Localization in the Mobile Ecosystem

- **Problem Statement:**
  The evolving Cross-Border Data Flows (CBDF)'s landscape is increasingly complex. Current state of data localization in CBDF, (i.e., data transfer as part of data governance), is fragmented due to divergent of policy, and regulatory frameworks. Some jurisdictions favour data localization based on national sovereignty, security and legitimate public policy. Some jurisdictions favour data localization based on sector specific legislative requirements, and some jurisdictions not in favour of data localization based on shared trust, accountability, responsibility and harmonized frameworks.

- **Motivation**:
  Mobile Network Operators (MNOs) process voluminous business process's dataset and consumer' datasets across the wider mobile and digital ecosystem. There are ongoing initiatives on Data Free Flow with Trust (DFFT) through use cases and technological solutions, greater legal certainty, complementary transfer mechanisms, and international regulatory co-operation on data privacy.

  GSMA source:
  Mobile Policy Handbook
  Cross-Border Data Flows: The impact of data localisation on IoT

  Relevant source:
  Moving Forward on Data Free Flow with Trust (DFFT)

- **Impact of data localization on mobile money services in Africa and Asia**
  In the context of mobile money, mobile money operators are facing several challenges: 1) Complex compliance procedures hampering the activities of mobile money operators; 2) Reduced efficiency in mobile money operations due to slow approval processes; 3) Constraints on the expansion of mobile money operators into other regions; 4) Limited access to the technology available to mobile money operators; 5) Increased cost of doing business; 6) Limits information sharing for fraud prevention and increases cybersecurity risks.

GSMA source:
Cross-border data flows: impact of data localisation on mobile money services in Africa and Asia

Forthcoming:
The GSMA Ministerial Programme 2025: Data Privacy Track

Tools and Practices

- Reputation Risk Management: Implement tools and practices that monitor potential reputational risks related to data sharing, ensuring transparency and addressing any concerns that could harm the collaboration's credibility.
- Privacy Oversight Systems: Engage internal privacy experts and establish systems to ensure data handling complies with relevant privacy regulations and standards.
- Interoperability Standards and Solutions: Ensure that data shared between organizations follows standardized formats and protocols, allowing

for smooth integration and compatibility across different systems.

- Risk Mitigation: Apply methods to assess and mitigate risks, such as privacy breaches or security issues, particularly when combining datasets that could lead to unintended consequences.

Resources for further reading

- Association of South-East Asian Nations (ASEAN), "Model Contractual Clauses (ASEAN MCCs) and Ibero-American Data Protection Network's Model Contractual Clauses (RIPD MCCs)" January, 2025.
- Digital Cooperation Organization (DCO), "Enabling Cross-Border Data Flows Amongst the Digital Cooperation Organization Member States." October, 2023.
- European Union, Model contract for data transfers
- ITU, Navigating Data Governance: A Guiding Tool for Regulators
- Singapore's Infocomm Media Development Agency (IMDA), Trusted Data Sharing Framework
- UNHCR. "PRIMES Interoperability Gateway (PING)"

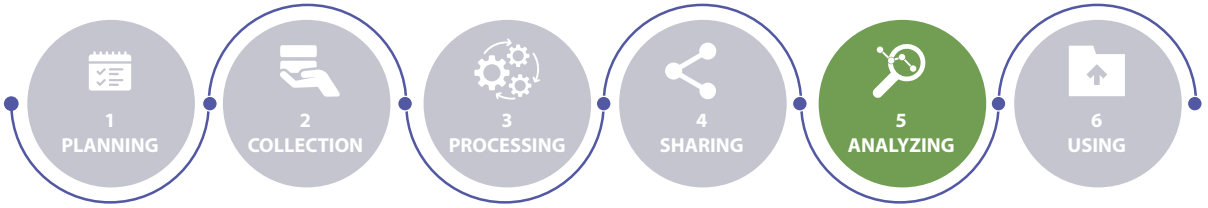
Box 8. Digital Public Infrastructure (DPI)

Digital public infrastructure (DPI) is an emerging approach to digital transformation that can improve service delivery at a societal scale. DPI is defined as a set of shared digital systems that are secure, interoperable, and built on open technologies, ensuring equitable access to public and private services at a societal scale. DPI utilizes common standards and reusable technology components. The core categories of DPI form the foundational structure to enable interoperability and reusability across systems and use-cases.

In this respect, the toolkit can help enhance DPI in the following ways:

- **Data privacy, security, and consent**: The toolkit promotes data privacy and ethical data use through recommending using data minimization and consent management tools. Similarly, in the context of DPI, this means government institutions usually can only access the necessary data, reducing risks and promoting trust. Privacy-by-design frameworks ensure that systems are set up with security and privacy from the outset, a foundational element for DPI. Consent management platforms can help ensure transparency in data collection and storage, which is crucial in DPI settings where public trust is essential.
- **Data quality and integrity**: The toolkit provides tools for data inventory, quality management, and bias detection, ensuring that data exchanged within DPI systems is relevant, accurate, and unbiased. Maintaining data integrity is key to DPI, where government and public sector entities rely on the same data sources for policy and service delivery.
- **Interoperability**: The toolkit advocates data sharing between organizations using standardized formats and protocols, allowing for smooth integration and compatibility across different systems. DPI systems are technically interoperable, allowing data to be efficiently accessed, used, and re-used across sectors.
- **Data storage and processing environments**: Secure storage and processing environments, access management, and data categorization practices create a robust infrastructure for DPI. This ensures data is not altered during processing and remains secure.

5. Analyzing



The analyzing stage is where data is interpreted to generate insights, inform decisions, and guide policy or programmatic action. It is the stage where the purposes and potential of data may be achieved—but also where risks of misuse, bias, or opaque decision-making may become apparent. Increasingly, data analysis includes AI and machine learning models, which require additional safeguards.

Given the sensitivity of certain types and methods of statistical analysis, it is vital to give consideration to factors such as outliers, robustness, algorithmic bias, sensitivity analysis and omitted variable bias (some types of analysis are sensitive to factors and data that have been left out of the data, as well as included in data samples). The choice of method(s) for processing and analysis has important consequences for the results, their robustness and sensitivity. Key governance considerations at this stage include aligning analysis with the original purpose(s), ensuring methodological transparency, testing for bias, and maintaining human oversight. Ethical implications should be continuously monitored, especially when insights may affect individual rights, public policy, or social outcomes.

Table 9. Checklist: Analyzing

Task	
Analysis plan aligned with original project goals	☐
Analytical methods, algorithms, and assumptions documented	☐
Bias testing conducted on data and models	☐
Interpretability and explainability mechanisms applied	☐
Outputs validated against real-world observations or ground truth	☐
Human oversight mechanisms in place for critical decisions	☐
Ethical review of analysis conducted (esp. for AI systems)	☐
Synthetic data or privacy-preserving methods used when needed	☐
Stakeholder engagement or participatory approaches considered	☐
Models retrained or recalibrated when significant changes in context occur	☐

10 Assessment Questions

1. Is the analysis aligned with the original stated purpose and problem definition?

Rationale: Purpose alignment prevents drift and ensures that results are relevant, actionable, and legally compliant.

2. Are the analytical methods, algorithms, and assumptions transparently documented?

Rationale: Transparency allows for replicability, auditability, and informed review by stakeholders.

3. Has the data and model been assessed for bias and representativeness?

Rationale: Bias in analysis can lead to discriminatory or inaccurate outcomes—testing helps detect and mitigate this risk.

4. Are the outputs interpretable and explainable, especially in the case of AI systems?

Rationale: Explainability builds trust and enables meaningful oversight of automated decisions.

5. Are analytical outputs validated against real-world outcomes or other benchmarks?

Rationale: Validation ensures accuracy, reliability, and guards against overfitting or false correlations.

6. Is human oversight embedded in the analysis-to-decision pipeline?

Rationale: Retaining human judgment is critical for ethical decision-making and error correction.

7. Have ethical risks (e.g., exclusion, surveillance, harm) been reviewed?

Rationale: Ethical review reduces harm and ensures the use of data is consistent with public values and human rights.

8. Are privacy-preserving technologies (e.g., synthetic data, PETs) used when analyzing sensitive data?

Rationale: These techniques allow meaningful analysis while protecting individuals' rights.

9. Are affected stakeholders informed or involved in interpreting and contextualizing results?

Rationale: Participatory analysis strengthens relevance and reduces the risk of harmful misinterpretation.

10. Are models regularly updated or retrained to reflect new data or changing conditions?

Rationale: Context evolves—updating models ensures continued relevance and accuracy over time.

Box 9. Artificial Intelligence (AI) and Data governance

As artificial intelligence (AI), including machine learning and generative AI, becomes an integral part of modern data analysis, new considerations arise in the data sharing and analysis stages. AI tools, which rely on large datasets to train models, bring both opportunities and risks.

- On one hand, AI can enable more efficient, nuanced, and scalable analysis, uncovering patterns and insights that were previously unattainable.
- However, the use of AI introduces concerns around data quality, bias, and interpretability.
- Machine learning models can unintentionally amplify biases present in training data, while generative AI may create outputs that are difficult to verify or interpret.
- Ensuring that these AI systems are transparent and ethical and that their decisions can be understood by humans is critical, especially where researchers are not sure why or how their models arrived at the results yielded.
- Additionally, AI-driven insights should be rigorously tested against real-world outcomes to prevent hidden variables or flawed predictions from influencing decision-making.
- Finally, human oversight remains crucial—while AI enhances analytical capabilities, retaining human control over the ultimate decisions ensures accountability and the ethical use of AI technologies in data governance.

Tools and Practices

- Data Modeling and Visualization: Design data structures optimized for targeted analysis. Ensure that data visualization tools translate complex data into clear, actionable insights.

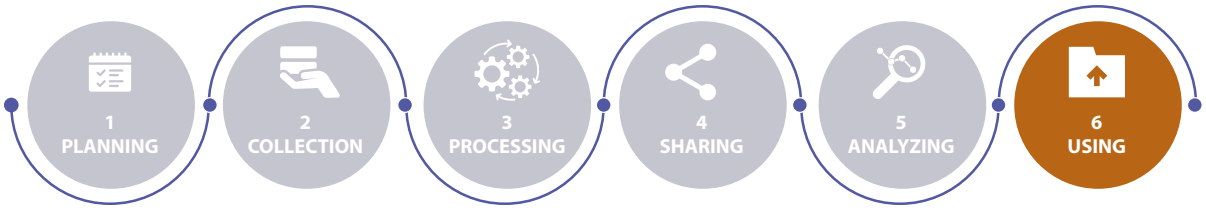
Resources

- UNDP Data to Policy Navigator, AI for Public Policy
- ASEAN Advisory Guidelines on Use Of Personal Data in AI Recommendation and Decision Systems
- ASEAN AI governance and ethics – generative AI

Use Cases

- UNDP Case Study on Ghana Marine Litter Project: AI can collect data faster and more accurately than humans. AI approaches can help improve the availability and the quality of data for addressing the SDG monitoring needs.
- UNDP Case Study on Gender Data Analysis: Breaking Barriers: Reinforcing Gender Data Analysis and Use with the Gender Data Lab Initiative

6. Using Data



The stage of using translates data-derived insights into action. This is where data and the resulting analysis can drive decisions, influence policies, or inform public services. It is also the stage where harm can be incurred or amplified if outputs are distorted, misunderstood, misapplied, or used beyond their intended purpose (e.g. via social media).

Good governance at this stage ensures that data use is consistent with original intentions, respects individual and collective rights, and includes processes for review, accountability, and learning. This includes final checks on consent, purpose alignment, and safeguards against misinterpretation or discriminatory application. Examples include peer review in academic journals, as well as the important scientific principle of verification and verifiability. As outputs are shared or acted upon, the importance of clear communication and continuous improvement becomes paramount. Indeed, how data are published and highlighted in different spheres can have major impact – consider, for example, the publication of COVID vaccination results and rates in medical journals versus the discussion around public perceptions of vaccination rates on social media.

Table 10. Checklist: Using

Task	
Final data use aligns with original project goals and consent	☐
Legal, ethical, and privacy compliance reviewed before use	☐
Risks of unintended consequences (e.g., re-identification, bias) assessed	☐
Findings reviewed by experts for quality and fairness	☐
Clear communication and messaging strategy developed	☐
Actions based on data are reviewed for effectiveness and impact	☐
Feedback mechanisms in place to capture user/community responses	☐
Data retention or secure deletion policies implemented	☐
Models retrained or revised based on new data or impact reviews	☐
Documentation of outcomes, learnings, and missed opportunities created	☐

10 Assessment Questions

1. Is the data being used in line with the original goals and consent agreements?

Rationale: Maintaining alignment prevents mission creep and ensures legal and ethical use of data.

2. Has the final use case undergone legal and ethical review?

Rationale: This confirms compliance and helps avoid unintended harms, particularly when data use impacts rights or services.

3. Are there processes to detect and mitigate potential unintended consequences?

Rationale: Anticipating downstream effects reduces the risk of public backlash or harm (e.g., from flawed policy decisions).

4. Have independent or multidisciplinary experts reviewed the findings or outputs?

Rationale: Peer review improves quality, credibility, and helps surface blind spots.

5. Is the data or insight being communicated clearly, with appropriate caveats?

Rationale: Clear communication prevents misinterpretation and supports informed public dialogue and policy decisions.

6. Are mechanisms in place to measure the actual impact of decisions informed by data?

Rationale: Evaluating outcomes ensures accountability and helps improve future data-driven efforts.

7. Is there a feedback loop from stakeholders or end-users on how the data was used?

Rationale: Feedback promotes responsiveness, inclusivity, and refinement of practices.

8. Are data retention and destruction policies followed post-use?

Rationale: Proper lifecycle closure supports privacy, legal compliance, and data minimization.

9. Are analytical models retrained or recalibrated over time?

Rationale: Keeping models up-to-date ensures continued relevance and avoids reliance on outdated insights.

10. Are lessons learned from the project documented for institutional memory?

Rationale: Capturing successes and failures supports learning, scaling, and policy adaptation.

Tools and Practices

- Comprehensive Review Systems: Implement review mechanisms to examine data quality, algorithmic outputs, and analysis integrity, avoiding errors or biases in reporting.
- Data Protection Protocols: Implement robust strategies to guard against re-identification risks and protect sensitive data, particularly during public release.
- Data Retention and Destruction Tools: Use data management systems to define long-term retention or secure deletion processes, ensuring compliance with privacy regulations such as the right to be forgotten.
- Model Maintenance and Retraining: Regularly update and retrain machine learning models with new data to keep analyses relevant and accurate over time.
- Reflection and Documentation Practices: Establish processes to evaluate and document project outcomes, focusing on both successes and missed opportunities for data use, to inform future initiatives.

Box 10. Cross Cutting Considerations to the Data Lifecycle

When managing data across its lifecycle, it is essential to consider cross-stage issues that support a seamless flow from collection to use.

- A well-structured data infrastructure is foundational, ensuring that each stage—whether collection, processing, sharing, or use—has the necessary technical and organizational support. This includes investing in secure data storage, interoperable platforms and protocols, and governance frameworks that enable data access and collaboration.
- Additionally, capacity building is critical across all stages. Stakeholders, including public servants, data stewards, and analysts, need continuous training to build technical expertise, data literacy, and awareness of ethical considerations.
- Equally important is fostering a data culture that promotes transparency, collaboration, and ethical data use. Establishing a culture of trust and responsibility, through change management practices, ensures that all actors across the lifecycle—from data collection teams to decision-makers—understand their role in protecting data integrity and using data to drive meaningful outcomes.

Glossary of Terms

Accountability
The obligation of decision-makers to ensure that data-related strategies, policies, and practices are participatory, transparent, and ethical. It requires those responsible for data governance to be answerable for their actions, fostering trust, fairness, and compliance with ethical and legal standards.

AI Governance
Frameworks, policies, and legal structures that shape the development, deployment, oversight, and societal effects of AI systems, ensuring they are ethical, transparent, accountable and uphold human rights and societal values. To effectively provide responsible AI systems, these frameworks should be rooted in data governance principles.[1]

Bias Mitigation in Data
Techniques and strategies to identify, reduce, or eliminate biases present in datasets and algorithms, ensuring that data-driven insights and decisions are fair and equitable.

Consent Management
A system or framework that enables individuals to give, review, and withdraw consent for the collection, processing, and sharing of their personal data, ensuring compliance with privacy regulations such as GDPR.

Cross-Border Data Governance
Policies, legal frameworks, and technical standards that regulate the movement, storage, and processing of data across national or regional jurisdictions.

Data Accountability Frameworks
Mechanisms and policies that ensure organizations and institutions take responsibility for how data is collected, managed, and used, often including auditing, reporting, and oversight structures.

Data Anonymization
The process of modifying personal data in a way that prevents the identification of individuals while still preserving the data's utility for analysis and decision-making.

Data Commons
A shared governance model that allows multiple stakeholders to contribute to, access, and use a collective pool of data resources while ensuring that ethical, legal, and social considerations are upheld.

Data Custodianship
The role of organizations or individuals responsible for maintaining and securing data assets on behalf of data owners or the broader public interest, ensuring compliance with governance frameworks.

Data Ecosystem
The integration of and interaction between different relevant stakeholders including data holders, data producers, data intermediaries and data subjects, that are involved in, or affected by, related data access and sharing arrangements, according to their different roles, responsibilities and rights, technologies, and business models.[2]

1 Verhulst S and Schüür F 'Interwoven Realms: Data Governance as the Bedrock for AI Governance' (Data & Policy Blog, 20 November 2023) https://medium.com/data-policy/interwoven-realms-data-governance-as-the-bedrock-for-ai-governance-ffd56a6a4543, accessed 10 February 2025.

2 OECD, Recommendation of the Council on Enhancing Access to and Sharing of Data (OECD Legal Instruments, 2019) https://legalinstruments.oecd.org/en/instruments/OECD-LEGAL-0463, accessed 10 February 2025.

Data Ethics
A new branch of ethics that studies and evaluates moral problems related to data (including generation, recording, curation, processing, dissemination, sharing and use), algorithms (including artificial intelligence, artificial agents, machine learning and robots) and corresponding practices (including responsible innovation, programming, hacking and professional codes), in order to formulate and support morally good solutions (e.g., right conducts or right values).[3]

Data Feminism
A perspective that highlights how power and privilege influence data collection, use, and governance, advocating for more responsive and equitable approaches to data-driven decision-making.

Data Intermediaries
Trusted third parties that facilitate the ethical and secure exchange of data between data providers and data users, ensuring compliance with governance frameworks while maximizing data utility.

Data Lifecycle
The set of processes in an application that transforms raw data into actionable knowledge.[4] These processes include collecting, processing, sharing, analyzing and using data governed by specific policies and practices at each phase.[5]

Data Minimization
The principle of 'data minimization' means that collection of personal information should be limited to what is directly relevant and necessary to accomplish a specified purpose.[6]

Data Mesh
An approach to data architecture where ownership of the data is distributed among cross-functional domain teams, who then provide data products to end users.[7]

Data Monetization
The process of generating economic value from data assets through licensing, partnerships, insights generation, or product development, often requiring governance mechanisms to ensure ethical and responsible use.

Data Openness Spectrum
A framework that categorizes data accessibility along a spectrum from closed (restricted access) to fully open (publicly available), with governance mechanisms defining appropriate levels of openness based on privacy, security, and ethical considerations.

Digital Public Infrastructure (DPI)
A set of shared digital systems that should be secure and interoperable and can be built on open standards and specifications to deliver and provide equitable access to public and/or private services at a societal scale and are governed by applicable legal frameworks and enabling rules to drive development, inclusion, innovation, trust, and competition and respect human rights and fundamental freedoms.[8]

---

3 Floridi L and Taddeo M, 'What Is Data Ethics?' (2016) 374 Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences 20160360, https://doi.org/10.1098/rsta.2016.0360.

4 NIST, 'Data Lifecycle' (NIST Glossary) https://csrc.nist.gov/glossary/term/data_life_cycle accessed 10 February 2025.

5 Young A, Zahuranec, A J, Verhulst, S and Gazaryan, K, *The Third Wave of Open Data Toolkit: Operational Guidance on Capturing the Institutional and Societal Value of Data Re-Use* (The GovLab 2021) https://files.thegovlab.org/The-Third-Wave-of-Open-Data-Toolkit.pdf accessed 10 February 2025.

6 European Data Protection Supervisor, Glossary (EDPS), https://www.edps.europa.eu/data-protection/data-protection/glossary/d_en#data_minimization accessed 10 February 2025.

7 Thoughtworks, 'Data Mesh' https://www.thoughtworks.com/en-de/insights/decoder/d/data-mesh accessed 10 February 2025.

8 UNDP, *Accelerating the SDGs through Digital Public Infrastructure: A Compendium of the Potential of Digital Public Infrastructure* (August 21, 2023 https://www.undp.org/publications/accelerating-sdgs-through-digital-public-infrastructure-compendium-potential-digital-public-infrastructure accessed 10 February 2025.

Data Provenance
The historical record of data, detailing its origins, transformations, and how it has been used over time to ensure accountability, reproducibility, and trust in data-driven decision-making.

Data Reciprocity
A principle that ensures fair and equitable data sharing among entities, emphasizing mutual benefit, transparency, and responsible stewardship in collaborative data governance arrangements.

Data Sandboxes
Controlled environments that allow for the safe testing and experimentation of data-driven innovations, balancing the need for regulatory oversight with the ability to explore novel applications of data.

Data Sensitivity Classification
The categorization of data based on its level of confidentiality, regulatory requirements, and potential impact on individuals or organizations if misused, helping guide appropriate governance measures.

Data Trusts
Legal and governance structures that enable collective stewardship of data, where trustees manage data assets on behalf of beneficiaries while ensuring compliance with ethical and legal frameworks

Digital Self-Determination (DSD)
The principle that individuals and communities have the autonomy to manage their digital presence, including control over personal data, digital identities, and the ways they engage with digital technologies. This concept emphasizes the empowerment of users to make informed decisions about how their data is collected, shared, and utilized, ensuring that digital interactions align with their values and interests. It encompasses both individual and collective dimensions, advocating for agency and rights throughout the digital data lifecycle.[9]

Data Sovereignty
A term frequently used interchangeably with digital sovereignty, cyber sovereignty, and technological sovereignty, this concept lacks a clear, widely accepted definition due to its geopolitical implications. Generally, it refers to the capacity of an entity—whether public or private—to exercise control over its digital future across three distinct layers. These layers include the physical layer (encompassing infrastructure and technology), the code layer (comprising standards, regulations, and design), and the data layer (pertaining to access, flow, and utilization of data).

Data Steward
Organizational leaders or teams empowered to create public value by re-using their organization's data (and data expertise); identifying opportunities for productive cross-sector collaboration and responding pro-actively to external requests for functional access to data, insights, or expertise. They are active in both the public and private sector, promoting trust within and outside their organization.[10]

Digital Identity Governance
The policies and standards that regulate how digital identities are created, authenticated, and managed to ensure security, privacy, and user control over personal data.

Differential Privacy
A privacy-enhancing technique that allows statistical analysis of datasets while ensuring that individual-level data remains undisclosed, reducing risks of re-identification.

Ethical AI Frameworks
Guidelines and governance structures that ensure AI systems are developed and deployed responsibly, addressing concerns such as bias, fairness, transparency, and accountability.

9 Verhulst, S, 'Operationalizing Digital Self-Determination' (2023) 5 Data & Policy e14 https://doi.org/10.1017/dap.2023.11

10 TheGovLab, 'Wanted: Data Stewards. (Re-)Defining the Roles and Responsibilities of Data Stewards for an Age of Data Collaboration' (2020) https://thegovlab.org/static/files/publications/wanted-data-stewards.pdf accessed 10 February 2025.

Federated Data Governance
A decentralized model of data governance where different entities maintain control over their respective data assets while following common standards and protocols for interoperability and ethical use.

Fair AI (Fairness, Accountability, and Transparency in AI)
A framework for evaluating AI systems to ensure they are equitable, unbiased, and explainable, aligning data governance practices with ethical principles.

Indigenous Data Sovereignty
Indigenous data sovereignty asserts the right of Indigenous communities to govern the collection, use, and stewardship of their data in accordance with their laws, customs, and knowledge systems. It challenges state-centered and market-driven models by advocating for self-determination, ethical data practices, and culturally appropriate governance structures.

Interoperability
The ability of two or more systems or applications to exchange information and to mutually use the information that has been exchanged.[11]

Knowledge Graphs in Data Governance
Structured representations of relationships between data elements that enhance interoperability, context-aware data integration, and responsible data sharing.

Metadata
Data about data.[12]

Privacy-by-Design
An approach that aims to protect individual privacy and data protection through intentional design choices. Unlike traditional privacy methods that view privacy as an afterthought, privacy by design makes privacy protection central starting from the very initial stages of design.[13]

Privacy-Enhancing Technologies (PETs)
Tools and methodologies that allow for data analysis while minimizing exposure to personally identifiable information (PII), including homomorphic encryption, secure multi-party computation, and synthetic data generation.

Proportionality in Data Governance
The principle that data governance measures should be proportionate to the level of risk, balancing data access and innovation with privacy, security, and ethical considerations.

Right to Data Portability
A legal right that allows individuals to obtain and transfer their personal data between service providers in a structured, commonly used, and machine-readable format.

Synthetic Data
Artificial data that is generated from original data and a model that is trained to reproduce the characteristics and structure of the original data.[14]

11 International Organization for Standardization (ISO), ISO/IEC 19941:2017 – Information Technology – Cloud Computing – Interoperability and Portability (ISO, 2017) https://www.iso.org/obp/ui/#iso:std:iso-iec:19941:ed-1:v1:en accessed 10 February 2025.

12 See (n. 2).

13 IEEE, Digital Privacy, 'What is Privacy-by-Design and why it is important?' https://digitalprivacy.ieee.org/publications/topics/what-is-privacy-by-design-and-why-it-s-important accessed 10 February 2025.

14 EDPS, Techsonar Report 2022-2023 (November 2022) https://www.edps.europa.eu/data-protection/our-work/publications/reports/2022-11-10-techsonar-report-2022-2023_en accessed 10 February 2025.

Transparency
In data governance, depending on context, transparency refers to one of the principles for processes and the principles that guide decision-making. As a principle of data governance processes, transparency ensures that governance processes are open and understandable. As a decision-making principle, transparency means clearly communicating the rationale behind decisions.

Algorithmic Transparency
The principle that decision-making processes powered by algorithms and AI should be explainable and understandable to those affected by them. This includes the ability to audit, interpret, and contest decisions made by automated systems.

Selected Data Governance Bibliography

1. **Data Governance Frameworks:** Documents on data governance models, principles, and international frameworks.

ADA Lovelace Institute. "Participatory data stewardship: A framework for involving people in the use of data." September 2021.

Barbero, Martina and Janet McLaren. "Effective and Ethical Data Sharing at Scale." Global Partnership for Sustainable Development Data. May 27, 2024.

Carnegie Endowment for International Peace. "Data Governance, Asian Alternatives: How India and Korea Are Creating New Models and Policies." Edited by Evan A. Feigenbaum and Michael R. Nelson. August 2022.

Centre for International Governance Innovation. "The Role of Governance in Unleashing the Value of Data." November 2024.

"Data Governance and Policy in Africa." Edited by Bitange Ndemo, Njuguna Ndung'u, Scholastica Odhiambo, Abebe Shimeles. Cham: Springer, 2023.

Datasphere Initiative. "Datasphere Governance Atlas 2022." April 2022.

European Commission. "A European Strategy for Data." February 19, 2020.

European Union. "The Digital Services Act." October 27, 2022.

Infocomm Media Development Authority. "Trusted Data Sharing Framework." July 1, 2024.

ITU. "Navigating Data Governance: A Guiding Tool for Regulators." October 31, 2024.

Lights on Data. "The complete guide to data governance roles and responsibilities."

MacFeely, Steve, Angela Me, Friederike Schueuer, Joseph M Costanzo, David Passarelli, Malarvizhi Veerappan, and Stefaan Verhulst. 2025. "Towards a Set of Universal Data Principles." *Statistical Journal of the IAOS: Journal of the International Association for Official Statistics* 41 (1): 150–55.

Marcucci, Sara, Natalia González Alarcón, Stefaan G. Verhulst, and Elena Wüllhorst. 2023. "Informing the Global Data Future: Benchmarking Data Governance Frameworks." *Data & Policy* 5 (January).

New Zealand Government. "Co-designing Māori data governance." 2021.

OECD. "Recommendation of the Council on Enhancing Access to and Sharing of Data." October 2021.

OECD. "Going Digital Toolkit"

Open Data Institute (ODI). "Mapping the Wide World of Data Sharing." July 4, 2019.

The Global Partnership on Artificial Intelligence (GPAI). "Data Governance Working Group: A Framework Paper for GPAI's Work on Data Governance 2.0." November 2022.

The GovLab. "Facilitating Data Flows through Data Collaboratives." October 19, 2023.

UNDP - Digital Strategy 2022-2025.

UNESCO. "Data sharing to foster information as a public good", 2023.

UNICEF. "UNICEF Data Quality Framework." April 2022.

United Nations Chief Executives Board (CEB). "Proposed Normative Foundations for International Data Governance: Goals and Principles." November 2024.

United Nations. "Global Digital Compact." September 22, 2024.

Verhulst, Stefaan G. "Operationalizing Digital Self-Determination." Data & Policy 5 (2023): e11. April 24, 2023..

World Economic Forum. "Advancing Digital Agency: The Power of Data Intermediaries." February 2022.

2. **Sector-Specific Frameworks:** Data governance models tailored to specific sectors, like health, humanitarian, and statistical data.

Data Responsibility Guidelines. OCHA, Centre for Humanitarian Data, 2025.

Dodgson, Kate, et al. A Framework for The Ethical Use of Advanced Data Science Methods in the Humanitarian Sector. Data Science and Ethics Group (DSEG), International Organization for Migration (IOM), 1 Apr. 2020.

GSMA. GSMA Guidelines on the Protection of Privacy in the Use of Mobile Phone Data for Responding to the Ebola Outbreak. GSMA, Oct. 2014.

Hastie, Rachel, and Amy O'Donnell. Responsible Data Management Training Pack. Edited by Sally Bolton, Oxfam, 2017.

International Organization for Migration. "DTM & Partners Toolkit: Enhancing Responsible Data Sharing." November 16, 2018.

Marelli, Massimo, et al., editors. Handbook on Data Protection in Humanitarian Action. Third Edition, Cambridge University Press, 2024.

OECD. "Recommendation of the Council on Health Data Governance." Adopted January 13, 2017.

Pan-American Health Organization (PAHO). "National Data Governance Framework: Information Systems for Health." October 25, 2024.

Sphere Association. The Sphere Handbook: Humanitarian Charter and Minimum Standards in Humanitarian Response. Fourth Edition, Sphere Association, 2018.

Transform Health. "Health Data Governance Principles." May 2021.

Verhulst, Stefaan. 2024. "The Need for Climate Data Stewardship: 10 Tensions and Reflections Regarding Climate Data Governance."

United Nations Statistics Division (UNSTATS). "United Nations Fundamental Principles of Official Statistics, Implementation Guidelines." January 2015.

IASC. "Operational Guidance on Data Responsibility in Humanitarian Action." April 2023.

WHO. "WHO Data Principles." June 2020.

3. **Data Protection, Privacy, and Security:** Sources covering data privacy, protection laws, and data security.

African Union "Convention on Cyber Security and Personal Data Protection" June, 2014

De France, James, and Luciia Laplante. IFRC Policy on the Protection of Personal Data. International Federation of Red Cross and Red Crescent Societies, 25 Mar. 2019.

European Union. "General Data Protection Regulation (GDPR)." April 27, 2016.

European Union. "Data Protection and Privacy Tools."

Information Commissioner's Office's Chapter 5: "Privacy-Enhancing Technologies (PETs) in Anonymisation and Pseudonymisation" September 2022.

IOM Data Protection Manual. International Organization for Migration, 2010.

ITU - International Telecommunication Union. "Technical Report D4.1 - Framework for security, privacy, risk and governance in data processing and management." December 2019.

OECD. "Emerging Privacy Enhancing Technologies: Current Regulatory and Policy Approaches." OECD Digital Economy Papers, March 2023.

OECD. "Recommendation of the Council concerning Guidelines Governing the Protection of Privacy and Transborder Flows of Personal Data." Adopted September 23, 1980; revised July 11, 2013.

Office of the Australian Information Commissioner. "Undertaking a Privacy Impact Assessment." May 2020.

Responsible Program Data Policy. Oxfam International, 17 Feb. 2025.

The Engine Room. "Hand-book of the Modern Development Specialist: Being, a Complete, Illustrated Guide to Responsible Data Usage, Manners, and General Deportment."

Task Team on Privacy-Enhancing Technologies of the United Nations Committee of Experts on Big Data and Data Science for Official Statistics. "UN Guide on Privacy-Enhancing Technologies for Official Statistics." 2023.

Verhulst, Stefaan. 2025. "Data Stewardship Decoded: Mapping Its Diverse Manifestations and Emerging Relevance at a Time of AI." SSRN Scholarly Paper.

UNICEF Policy on Personal Data Protection. UNICEF, 15 July 2020.

UN Development Group (UNDG). "Data Privacy, Ethics and Protection: Guidance Note on Big Data for Achievement of the 2030 Agenda." November 2017.

UN Global Pulse. "Risks, Harms and Benefits Assessment." June 2018.

UNESCO. "Freedom of Information: A comparative legal survey." 2008.

UNESCO. "Human Rights and Encryption." 2016.

UNESCO. "Guidelines for judicial actors on privacy and data protection." 2022.UNHCR. "Policy on the Protection of Personal Data of Persons of Concern to UNHCR." May 2015.

UNHCR , General Policy on Personal Data Protection and Privacy, 2022

UN High Commissioner for Refugees (UNHCR), Policy on the Protection of Personal Data of Persons of Concern to UNHCR, May 2015.

4. **Data Power, Justice, Ethics, and Digital Rights:** Works exploring the ethics, justice, and social implications of data, including power dynamics and individual rights.

Ada Lovelace Institute. "Rethinking Data and Rebalancing Digital Power." November 2022.

Access Now and Government of Catalonia. "The Geneva Declaration on Targeted Surveillance and Human Rights." September 29, 2022.

Center for Human Rights and Global Justice (CHRGJ). "Paving a Digital Road to Hell? A Primer on the Role of the World Bank and Global Networks in Promoting Digital ID." June 2022.

D'Ignazio, Catherine, and Lauren F. Klein. Data Feminism. The MIT Press, March 2020; reprint October 2023.

Data Ethics Commission (Germany). "Opinion of the Data Ethics Commission." October 2019.

Global Partnership on Artificial Intelligence (GPAI). "Data Justice: A Primer on Data and Economic Justice." November 2022.

Heeks, Richard, and Jaco Renken. "Data Justice for Development: What Would It Mean?" Information Development, vol. 34, no. 1, 2018, pp. 90–102.

International Organization for Migration (IOM). "IOM Data Protection Manual." 2010.

Open Data Institute (ODI). "The Data Ethics Canvas." Last updated June 2021.

Taylor, Linnet. "What is Data Justice? The Case for Connecting Digital Rights and Freedoms Globally." Big Data & Society, vol. 4, no. 2, July–December 2017.

Verhulst, Stefaan G. "Reimagining data responsibility: 10 new approaches toward a culture of trust in re-using data to address critical public needs." Data & Policy 3 (2021)

Zahuranec, Andrew J., Hannah Chafetz, and Stefaan Verhulst. 2023. "Data Sharing Agreements:

Moving from Idea to Practice." Brooklyn, New York.

5. **Children's Digital Rights:** Documents and guidelines specifically addressing children's rights in digital spaces.

Council of Europe. "Guidelines to Respect, Protect, and Fulfil the Rights of the Child in the Digital Environment." Adopted July 4, 2018.

Global Privacy Assembly. "Resolution on Children's Digital Rights." October 25, 2021.

"Principles." Responsible Data for Children.

UNICEF and The GovLab. "Responsible Data for Children Synthesis report."

UNICEF Office of Research-Innocenti. "The Case for Better Governance of Children's Data: A Manifesto." 2020. Accessed October 25, 2024.

United Nations. "Children's Rights in the Digital Environment." December 19, 2023.

United Nations. "Guidance Note of the Secretary-General, Child Rights Mainstreaming." July 2023.

United Nations. "Promotion and protection of the rights of children." December 2023.

5 Rights. "Updated Disrupted Childhood: The cost of persuasive design." April 11 2023.

5 Rights. "Risk by Design."

UN Committee on the Rights of the Child. "General comment No. 25 (2021) on children's rights in relation to the digital environment." 2021.

6. **Data for Development:** Data-driven approaches supporting sustainable development.

AI in Global Development Playbook. USAID, DOS, 2024.

African Union. "The Digital Transformation Strategy for Africa (2020–2030)." 2020.

OECD. "Measuring the Value of Data and Data Flows ."

Paul, Amy, et al. Reflecting the Past, Shaping the Future: Making AI Work for International Development. Center for Digital Development at USAID, 2018.

Paul, Amy, et al. Managing Machine Learning Projects in International Development. USAID, DAI, Vital Wave, 2021.

Verhulst, Stefaan, Laura Sandor, Elena Murray, and Peter Addo. 2024. "Responsible Data Re-Use in Developing Countries: Social Licence through Public Engagement."

UNDP. "Accelerating The SDGs Through Digital Public Infrastructure: A Compendium of the Potential of Digital Public Infrastructure." Published August 21, 2023.

UNDP. "Data Principles for UNDP." 2020.

Verhulst, Stefaan "Social License for Data: Going Beyond Consent to Protect Vulnerable Populations"

World Bank. "World Development Report 2021: Data for Better Lives." Published March 2021.

Verhulst, Stefaan G. "Reusing data responsibly to achieve development goals." In OECD *Development Co-operation Report 2021: Shaping a Just Digital Transformation*, 289-297. Paris: OECD Publishing, 2021

World Bank. "ID4D Practitioner's Guide."

7. **Cross-Border Data Flows and Digital Economy**

Brookings. "Data portability and interoperability: A primer on two policy tools for regulation of digitized industries." May, 2023.

Catapult. "City Data Sharing Toolkit." September 2019.

Datasphere Initiative. "Sandboxes for Data: Creating Spaces for Agile Solutions Across Borders." May 2022.

Digital Cooperation Organization. "Enabling Cross-Border Data Flows amongst the Digital Cooperation Organization Member States." October 2023.

Digital Public Goods Charter. "Charter for Digital Public Goods." June 1, 2022.

Gebru, Timnit et al. "Datasheets for Datasets." December 1, 2021.

The GovLab "Facilitating Data Flows through Data Collaboratives"

Global Data Alliance. "Cross-Border Data Policy Principles." March 2, 2021.

Global Partnership for Sustainable Development Data. "Data Sharing Cookbook." May 2024.

NESTA, UK Government "Data Sharing Toolkit"

Open Data Institute. "Mapping the Wide World of Data Sharing."

Verhulst, Stefaan. 2023. "Policy Paper 10 – Data Collaboratives: Enabling a Healthy Data Economy through Partnerships." In The Digital Revolution and the New Social Contract. IE University

UNCTAD. "Digital Economy Reports."

8.AI Governance and Regulation: Regulations, principles, and ethical guidelines for AI development and application.

Adams, Rachel, et al. Global Index on Responsible AI. Global Center on AI Governance, 2024.

AI Act Explorer. https://artificialintelligenceact.eu/ai-act-explorer/

Artificial Intelligence Plan: Charting the Course for Responsible AI in USAID Programming. USAID, 2022.

European Commission. (2024). *Artificial Intelligence Act (Regulation (EU) 2024/1689)*. https://eur-lex.europa.eu/eli/reg/2024/1689/oj/eng

The GovLab. "Resources on Data Sharing Agreements."

Global Partnership in Artificial Intelligence (GPAI). "Global Index on Responsible AI."

Hugging Face. "Building Better AI: The Importance of Data Quality."

OECD. "OECD AI principles." Adopted May 22, 2019.

OECD. "Regulatory Sandboxes in Artificial Intelligence." May 2022.

"The IEEE Global Initiative 2.0 on Ethics of Autonomous and Intelligent Systems." IEEE Standards Association.

UN System Chief Executives Board for Coordination (CEB): Principles for the Ethical use of AI by the UN System.

UNESCO. "UNESCO Recommendation on the Ethics of AI." November 23, 2021.

9.Interoperability

European Commission. "ISA^2 - Interoperability solutions for public administrations, businesses and citizens." April 25, 2018.

European Union. "Access to Base Registries: Good Practices on Building Successful Interconnections of Base Registries." 2016.

European Union: Good Practices on Building Successful Interconnections of Base Registries

UNHCR. "PRIMES Interoperability Gateway (PING)"

10. Data Governance for Business
Resources focusing on the application of data governance principles and strategies within business contexts.

Aubert, Benoit, et al. Data Governance: Governing Data for Sustainable Business. Edited by Alison Holt, BCS, The Chartered Institute for IT, 2020.

Bhansali, Neera. Data Governance: Creating Value from Information Assets. CRC Press, Taylor & Francis Group, 2014.

Deloitte. "Data Valuation: Understanding the Value of Your Data Assets." 2020.

Eryurek, Evren, et al. Data Governance: The Definitive Guide: People, Processes, and Tools to Operationalize Data Trustworthiness. 1st edition, O'Reilly Media, 2021.

Singapore Infocomm and Media Development Authority (IMDA) and Personal Data Protection

Commission (PDPC). "Guide to Data Valuation for Data Sharing." 2019.

International, DAMA. The DAMA Guide to the Data Management Body of Knowledge - Print Edition. First Edition, Technics Publications, LLC, 2010.

Ladley, John. Data Governance: How to Design, Deploy and Sustain an Effective Data Governance Program. Newnes, 2012.

Madsen, Laura B. Disrupting Data Governance: A Call to Action. Technics Publications, 2019.

Maydanchik, Arkady. Data Quality Assessment. Technics Publications, LLC, 2012.

OECD. "Measuring the Value of Data and Data Flows." OECD Digital Economy Papers. Paris:

OECD Publishing, December 14, 2022.

Reichental, Jonathan. Data Governance For Dummies. 1st edition, For Dummies, 2022.

Appendix: Checklists

WHY: Determining Purposes for Data and Data Governance

Please check all purposes that apply to your data governance initiative:

Maximizing Data Utility and Value / Minimizing Harms and Costs	
Unlock data for new insights and improved decision-making.	☐
Create tangible value for the organization and/or society through data use.	☐
Minimize potential harms, costs, and unintended negative consequences associated with data collection and use.	☐
Fostering Innovation and Sustainable Development	
Stimulate innovation and entrepreneurship using data-driven approaches.	☐
Promote and enhance economic opportunities through data.	☐
Advance Sustainable Development Goals (SDGs) using data for monitoring and implementation (e.g., in health, education, infrastructure).	☐
Foster a data-driven culture and enhance data skills within the organization/society.	☐
Establishing Equity and Digital Self-Determination	
Promote equitable access to data and ensure benefits are shared fairly.	☐
Ensure marginalized or vulnerable groups, including children, are actively included and protected in data ecosystems.	☐
Uphold principles of data self-determination for individuals and communities.	☐
Implement specific governance frameworks to protect the data of vulnerable populations.	☐
Supporting Specific Policy or Operational Objectives	
Enhance transparency, accountability, and citizen engagement (e.g., through Open Data initiatives).	☐
Support national SDG priorities such as accommodating the impact of climate change or advancing education for all.	☐
Facilitate safe and effective cross-border data sharing and international collaboration while ensuring legal compliance.	☐
Mobilize data effectively for crisis preparedness, risk management, response, and recovery efforts.	☐
Support the responsible development and deployment of Artificial Intelligence (AI), addressing bias, fairness, and ethical concerns.	☐
Align with or implement Digital Public Infrastructure (DPI) principles to enable secure, private, interoperable, and seamless data exchange.	☐

HOW: Determining Data Governance Principles

This checklist provides a way to determine and evaluate the presence of principles across three categories:

1. **Principles for Processes** – how governance decisions are made
2. **Principles for Decisions** – what informs those decisions
3. **Principles for Data Handling** – how data is managed in practice

Principles for Processes	
These principles ensure governance activities are conducted in a fair, ethical, and transparent way.	
Transparency – Governance processes are open and understandable.	☐
Accountability – Actors are responsible for outcomes of governance processes.	☐
People-Centeredness – The needs and rights of individuals are prioritized.	☐
Fairness – Equal and unbiased treatment in procedures.	☐
Participation – Stakeholders are meaningfully involved in governance.	☐
Lawfulness – All actions comply with relevant laws and regulations.	☐
Inclusiveness – Marginalized groups and those affected by datafication are represented.	☐
Principles for Decisions	
These principles guide how decisions in data governance are defined and implemented.	☐
Transparency – The rationale behind decisions is clearly communicated.	☐
Proportionality – Decisions are appropriate to their context and impact.	☐
Defined Purpose – Decisions are guided by clear and specific objectives.	☐
Accountability – Decision-makers are answerable for their choices.	☐
People-Centeredness – Individuals' needs and rights are front and center.	☐
Fairness – Decisions are just and equitable.	☐
Protection from Harm and Non-Discrimination – Risks are mitigated, and bias is avoided.	☐
Participation – Diverse perspectives are integrated into decision-making.	☐
Principles for Data Handling	
These principles ensure that data is managed responsibly, securely, and in line with rights and expectations.	
Confidentiality and Security – Sensitive data is protected from unauthorized access.	☐

Proportionality – Data practices are aligned with the purpose and necessity.	☐
Accessibility and Portability – Data is available and portable under appropriate conditions.	☐
Protection of Privacy – Personal data is safeguarded according to privacy laws.	☐
Lawfulness – Data handling complies with all applicable legal standards.	☐
Informed Consent – Data subjects are aware of and agree to data use.	☐
Data and Metadata Quality – Data is accurate, reliable, and well-documented for reuse.	☐
Interoperability and Standardization – Data adheres to shared standards for easier exchange and integration.	☐

WHO: Determining Roles and Responsibilities

RACI	Planning	Collection	Processing	Sharing	Analyzing	Using
Responsible						
Accountable						
Consulted						
Informed						

WHAT: Planning

Task	
Purpose and value of data clearly defined and documented	☐
Stakeholders and affected communities identified and mapped	☐
Legal and policy landscape reviewed, including prior efforts	☐
Governance model designed (roles, responsibilities, decision-making)	☐
Scope, goals, and limitations of the data project articulated	☐
Financial, technical, and human resources evaluated and secured	☐
Data sharing agreements (MoUs, DSAs) and legal templates prepared	☐
Interoperability needs and shared vocabularies discussed	☐
Risk assessments conducted (e.g., privacy, security, AI risks)	☐
Stakeholder engagement and trust-building strategy in place	☐
Communication and transparency plan drafted	☐
Feedback loops and evaluation metrics defined	☐

WHAT: Collection

Task	
Data collection aligns with stated purpose and minimization principles	☐
Consent mechanisms established and documented (including dynamic consent)	☐
Marginalized or underserved populations are represented	☐
Privacy-by-design principles applied in collection systems	☐
Data formats standardized and interoperable	☐
Metadata and documentation practices defined	☐
Cross-border data flow and localization requirements assessed	☐
Tools in place to detect and mitigate bias during collection	☐
Encryption or anonymization methods applied at point of collection	☐
Legal obligations reviewed and adhered to (e.g., lawful basis, jurisdictional issues)	☐
Feedback mechanisms in place to adapt collection processes if issues arise	☐

WHAT: Processing

Task	
Data cleaning and transformation processes documented	☐
Data quality (accuracy, consistency, completeness) validated	☐
Data categorized and classified (e.g., sensitive vs. non-sensitive)	☐
Data provenance and version history maintained	☐
Encryption and access controls implemented	☐
Internal access protocols defined (e.g., tiered access, audit logs)	☐
Privacy-enhancing technologies considered or applied	☐
Backup, archival, and secure deletion systems in place	☐
Compatibility with interoperability standards ensured	☐
Processing activities reviewed for legal and ethical compliance	☐
Mechanisms in place to avoid unintended data linkages	☐

WHAT: Sharing

Task	
Purpose and scope of data sharing clearly defined and agreed	☐
Legal agreements (e.g., DSAs, MoUs, licensing) formalized	☐
Internal stakeholders aligned and briefed on governance expectations	☐
External communication strategy developed	☐
Reputational risks and mitigation plans identified	☐
Interoperability protocols established (e.g., common vocabularies, formats)	☐
Technical safeguards in place (e.g., secure APIs, federated access)	☐
Privacy and security risks from data combination anticipated	☐
Governance for third-party intermediaries or platforms clarified	☐
Social license and ethical considerations reviewed	☐
Ongoing coordination and review mechanisms established	☐

WHAT: Analysis

Task	
Analysis plan aligned with original project goals	☐
Analytical methods, algorithms, and assumptions documented	☐
Bias testing conducted on data and models	☐
Interpretability and explainability mechanisms applied	☐
Outputs validated against real-world observations or ground truth	☐
Human oversight mechanisms in place for critical decisions	☐
Ethical review of analysis conducted (esp. for AI systems)	☐
Synthetic data or privacy-preserving methods used when needed	☐
Stakeholder engagement or participatory approaches considered	☐
Models retrained or recalibrated when significant changes in context occur	☐

WHAT: Using

Task	
Final data use aligns with original project goals and consent	☐
Legal, ethical, and privacy compliance reviewed before use	☐
Risks of unintended consequences (e.g., re-identification, bias) assessed	☐
Findings reviewed by experts for quality and fairness	☐
Clear communication and messaging strategy developed	☐
Actions based on data are reviewed for effectiveness and impact	☐
Feedback mechanisms in place to capture user/community responses	☐
Data retention or secure deletion policies implemented	☐
Models retrained or revised based on new data or impact reviews	☐
Documentation of outcomes, learnings, and missed opportunities created	☐

BROADBAND
COMMISSION
FOR SUSTAINABLE
DEVELOPMENT

broadbandcommission.org

Photo cover: metamorworks/Shutterstock.com