

RESEARCH ARTICLE

# Informing the Global Data Future: Benchmarking Data Governance Frameworks

Sara Marcucci<sup>1</sup>, Natalia González Alarcón<sup>1</sup>, Stefaan G. Verhulst<sup>1,2,3,4,5</sup>  and Elena Wüllhorst<sup>1</sup>

<sup>1</sup>The GovLab, Brooklyn, NY, USA

<sup>2</sup>Center for Urban Science and Progress, New York University, New York, NY, USA

<sup>3</sup>ISI Foundation, Turin, Italy

<sup>4</sup>Vrije Universiteit Brussel, Brussels, Belgium

<sup>5</sup>The Data Tank, Brussels, Belgium

**Corresponding author:** Stefaan G. Verhulst; Email: [stefaan@thegovlab.org](mailto:stefaan@thegovlab.org)

**Received:** 13 April 2023; **Revised:** 04 July 2023; **Accepted:** 05 July 2023

**Keywords:** data governance; data protection; international development; policy; responsible data use

## Abstract

Data has become a critical trans-national and cross-border resource. Yet, the lack of a well-defined approach to using it poses challenges to harnessing its value. This article explores the increasing importance of global data governance due to the rapid growth of data, and the need for responsible data practices. The purpose of this paper is to compare approaches and identify patterns in the emergent data governance ecosystem within sectors close to the international development field, ultimately presenting key takeaways and reflections on when and why a global data governance framework may be needed. Overall, the paper provides information about the conditions when a more holistic, coordinated transnational approach to data governance may be needed to responsibly manage the global flow of data. The report does this by (a) considering conditions specified by the literature that may be conducive to global data governance, and (b) analyzing and comparing existing frameworks, specifically investigating six key elements: purpose, principles, anchoring documents, data description and lifecycle, processes, and practices. The article closes with a series of final recommendations, which include adopting a broader concept of data stewardship to reconcile data protection and promotion, focusing on responsible reuse of data to unlock socioeconomic value, harmonizing meanings to operationalize principles, incorporating global human rights frameworks to provide common North Stars, unifying key definitions of data, adopting a data lifecycle approach, incorporating participatory processes and collective agency, investing in new professions with specific roles, improving accountability through oversight and compliance mechanisms, and translating recommendations into practical tools.

## Policy Significance Statement

This research article highlights the need for a coordinated transnational approach to data governance to ensure responsible data promotion and protection for the public interest. The research compares approaches and identifies patterns in the emergent and fragmented data governance ecosystem within sectors close to the international development field, ultimately presenting a series of reflections on when and why a global data governance framework may be needed. The findings presented in this report can inform policymakers, regulators, and relevant stakeholders to work toward a global consensus on data governance, leading to the development of a common framework that can enhance the value of data and promote its responsible use.

## 1. Introduction: Identifying Patterns in a Fragmented Data Governance Ecology

Data has become a global asset; therefore, how it is managed and governed has become a priority for a diverse number of stakeholders around the world. Historically, data governance has often been associated with private organizational or corporate governance approaches. However, as technological innovation and the amount of data have increased rapidly in recent years, the notion of data governance has evolved to include governmental and institutional bodies. There have been increasing calls for a global data governance framework that would help manage the global flow of data responsibly, while ensuring a necessary balance between its undeniable potential and equally undeniable risks. The World Development Report 2021 by the World Bank, for instance, acknowledges the increasing development of data governance arrangements, yet incipient, and alerts on how the current regulatory efforts might be inadequate for the “majority world” (Mungai et al., 2022). There are different reasons associated, some related to significant gaps in infrastructure, security, institutional and safeguard mechanisms, and others linked to countries’ unlike needs and priorities. “A global consensus would give individuals and enterprises confidence that data relevant to them carry similar protections and obligations no matter where they are collected or used. (...) It would also establish ground rules for the exchange of data between commercial use and the public good” (World Bank, 2021, 297).

The search for a global data governance framework emerges from a complex landscape that bridges policy and practice, and encompasses a number of different domains, such as data management, data ethics, and data protection. “The approach to governing data and data flows varies considerably among the major players in the digital economy, and there is little consensus at the international and regional levels” (UNCTAD, 2021, 98). In addition, different frameworks within and across countries, regions, sectors, and organizations have resulted in a patchwork of policies, frameworks, and practices, leading to a fragmented ecology that poses certain challenges to the evolution of a common framework.

This fragmentation is heightened by the dynamism of the ecology, with new solutions, often technical in nature, being released on a regular basis, often seeking to update legacy approaches that are no longer fit for the challenges and opportunities of new data realities. In addition, emerging technologies and concepts, for instance, artificial intelligence or distributed ledger technologies, lead to new governance needs and result in sometimes ad-hoc extensions to existing data governance approaches (World Bank, 2021, 271; Anthony, 2022, 293).

While some of the frameworks—such as data protection regimes—are more mature and lend themselves to standardization and codification, others fall short, representing more reactive approaches. As a result, efforts to harmonize and coordinate the various frameworks have often been led by business or professional associations, standard-setting bodies, or international fora connecting national data protection authorities rather than by any entity with truly global reach and credibility. Because of the diversity of entry points, concerns, and interests, there is a wide diversity of actors advocating for different approaches, often operating in silos without much engagement or coordination. However, the increasing interconnection and interdependence within the global data economy urge to evolve toward a more holistic, coordinated transnational approach that might require new and innovative global governance (UNCTAD, 2021, 215).

This report aims to examine and analyze the evolving and fragmented data governance landscape in sectors related to international development. By considering conditions specified in the literature that favor global data governance and conducting a comprehensive analysis of existing frameworks, the report provides insights into the reflections on the necessity of a global data governance framework, highlighting the conditions and reasons that may warrant such an approach.

## 2. Why a Global Governance Framework May Be Necessary

This section aims to summarize some of the existing literature exploring why and when global governance may be beneficial, focusing particularly on (a) global coordination to prevent harmful fragmentation,

(b) the advancement of global principles and values, and (c) using data as a resource to advance global public goods.

### ***2.1. Global coordination to prevent harmful fragmentation***

As the world becomes increasingly interconnected, it is ever more urgent to build systems of cooperation that allow multiple and diverse actors to collaborate and make use of a dedicated framework for sharing information, expertise, and experience. It could seem particularly crucial to develop a standardized approach to data in certain sectors or at certain moments, such as in times of humanitarian crisis, as this will ultimately enable more coordination and prevent potentially harmful fragmentation. Given the multiple and complex interactions between regulations and asymmetries at local, national, and international levels, fragmentation may have profound implications on individuals and businesses, both intended and unintended, for virtually all aspects of our daily lives (Fay, 2022).

Further, the absence of a systematic global approach to data governance may create inequitable consequences for low- and middle-income countries as it would be harder for them to participate in the global digital economy and develop their own frameworks responsibly (Pisa and Nwankwo, 2021). Even more ambitious approaches argue that a “new Bretton Woods-style agreement” is necessary to redefine the global governance model in a digital and hyper-globalized world (Medhora and Owen, 2020). By enhancing cooperation and shared standards and principles, global data governance might allow nations and organizations to collectively take advantage of the potential data harbors to face common challenges and respond to collective needs.

### ***2.2. Advancing global principles and values***

A global data governance framework would enable international cooperation and coordination to promote globally shared principles and values, such as human rights, and to further anchoring frameworks such as the Universal Declaration of Human Rights or the SDGs. In fact, data are playing an increasingly important role in the humanitarian field. From the United Nations to private non-governmental organizations (NGOs), organizations across the world are starting to adopt data on ever larger scales to enable more agile, efficient and evidence-based decision-making to promote human rights and other global values.

For example, the World Economic Forum (WEF) has proposed a new data governance model called Authorized Public Purpose Access (APPA), defined as “a model for realizing value by permitting access to data for specific, agreed public purposes (...) through processes that do not rely exclusively on explicit, individual consent as a means of protecting human rights” (World Economic Forum (WEF), 2021). The ultimate goal is guaranteeing individual human rights regarding data use, not limited to privacy rights. Also, as part of the 2030 Agenda, the UN reaffirmed its commitment to international law and emphasized that all efforts shall be implemented in a manner that is consistent with human rights (United Nations High Commissioner for Human Rights, 2018). Given the international and often borderless nature of these objectives, and considering the key role data has in achieving them, a global data governance framework is essential.

### ***2.3. Using data as a common resource to advance global public goods***

Increasingly, data are a resource for supporting global public goods. It is crucial that no single entity has sovereignty over data and that these collective public goods are managed responsibly, in a manner that secures and preserves them for all of humankind. Lately, there has been a debate on treating data as “commons” when the economic characteristics of data define it as an intangible non-rival asset, which may suggest an opposite definition. However, it is a practical approach when considering data governance and regulation.

Ostrom’s principles (Ostrom, 2012) for managing shared resources (commons) offer insightful guidelines to reach an agreement about rules for accessing data when some individuals may need to

sacrifice personal benefit for the greater common good (Coyle et al., 2020). The principles help understand the asymmetries of information and incomplete agreements that characterize the data economy and provide innovative ways to govern shared resources. For example, defining the rights of different entities to control, access, use and share data, and establishing monitoring and auditing mechanisms for data use and sharing (Coyle et al., 2020) might help to address the governance challenge of preventing misuse of sensitive data while fostering the reuse of data to create social value and potential “knowledge spillovers” (Coyle, 2020; Aaronson, 2022).

### 3. Methodology

The findings and recommendations of this article are based both on empirical data as well as an analytical framework, which permits to derive broader conclusions from the data.

#### 3.1. Sampling strategy

In conducting this research, more than 100 data governance documents were identified, of which 58 were curated and surveyed in detail, across 37 organizations (non-governmental, intergovernmental, and independent), 8 national or local government entities, and 4 regional bodies. [Appendix A](#) of the Supplementary Material includes the full list of the frameworks analyzed in depth, while the Template of Analysis in Section 3.3 provides a detailed assessment strategy of the frameworks considered. In addition, [Appendix B](#) includes the link to the public repository of cases.

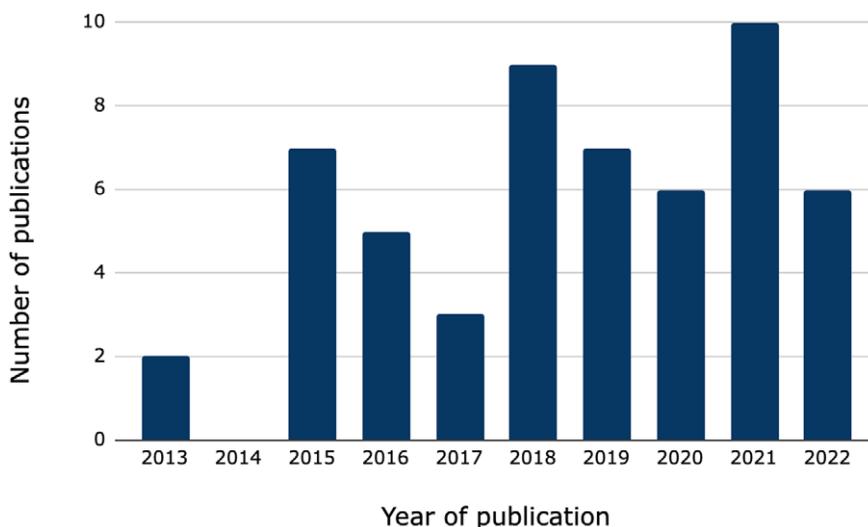
This article was first initiated as a report to inform the United Nations High-Level Committee on Programmes (HLCP) and has subsequently evolved to provide insights to other intergovernmental and governmental institutions, NGOs, academic and research institutions, and other stakeholders that aim to use and govern data.

The sampling strategy took into account the following considerations:

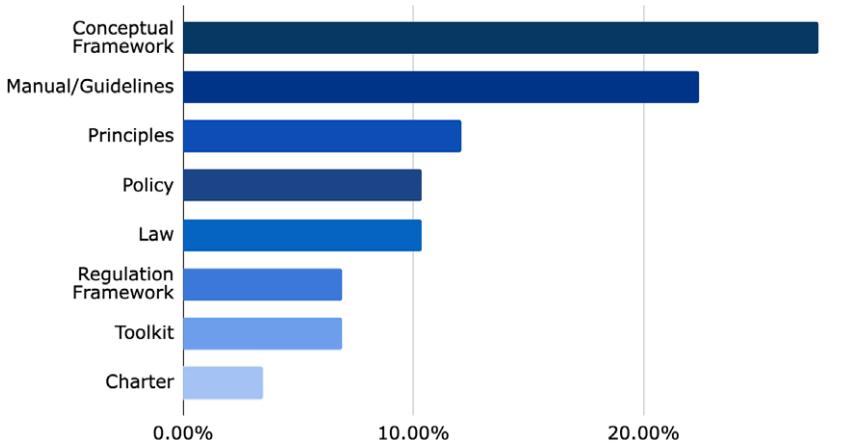
*Timeframe:* To ensure relevance and validity, this research focused on frameworks created within the last 10 years (2013–2022; [Figure 1](#)).

*Variety of frameworks:* This research aims to represent the variety and heterogeneity of existing data governance frameworks and approaches, and thus includes a wide range of examples and formats, encompassing principles, guidelines, and regulatory frameworks and standards ([Figure 2](#)).

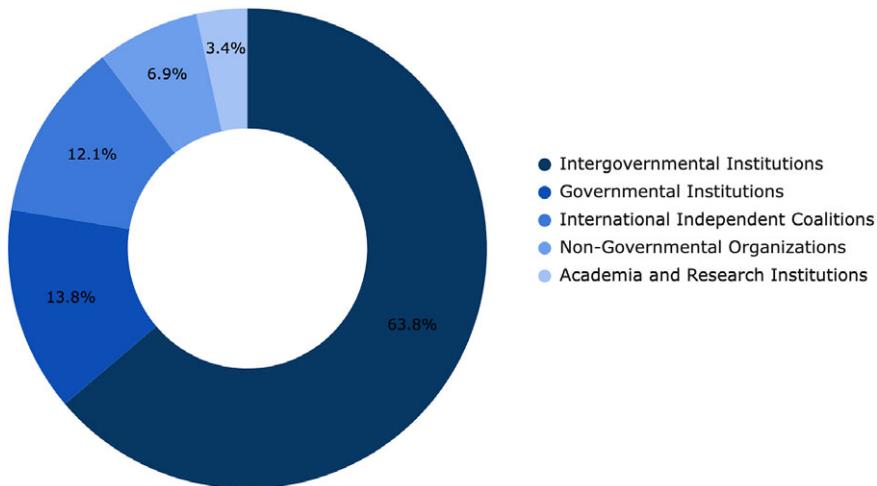
*Types of organizations:* The research prioritized public sector efforts over private-sector-oriented ones, cognizant of how the public sector plays a critical role in setting policies and regulations for data



**Figure 1.** Number of publications per year.



*Figure 2. Types of frameworks.*



*Figure 3. Type of organization.*

governance. While private sector efforts are also important in this regard, to maintain the scope of the research manageable, they were not the primary focus of this research. Thus, the sample prioritizes governmental and intergovernmental institutions, NGOs, academic and research institutions, and international independent (sectoral) coalitions (Figure 3).

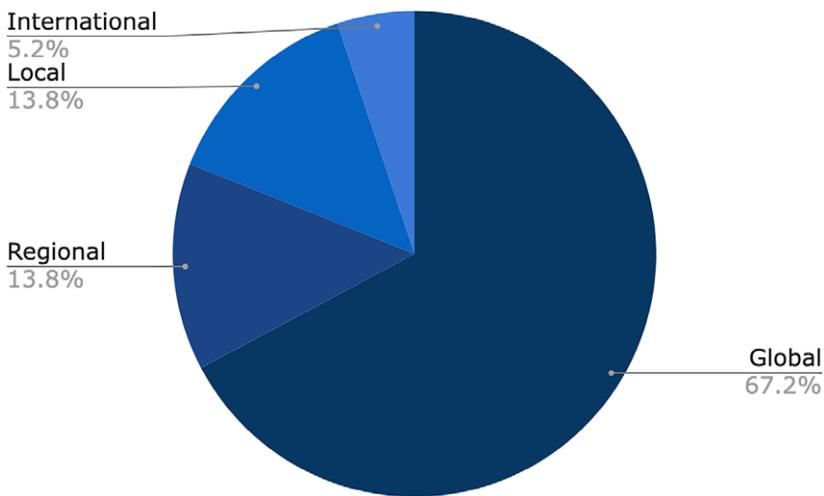
*Geographical scope:* When considering the geographical scope, this research aims to cover various levels of jurisdictions, including global, regional, and national frameworks. Although diversity and comprehensive representation across regions was sought, a stocktaking of various data governance frameworks and practices from the Global South is still largely missing. On the one hand, that is due to a lack of a clear and widely spread understanding of the frontier of data governance best practices. This might limit some governments from finding a suitable reference when building and prioritizing their own (Chen, 2021). On the other hand, sometimes legal and regulatory frameworks for data are inadequate in lower-income countries, which too often face substantial gaps in safeguards and shortages of infrastructure. Indeed, as the World Development Report 2021 by the World Bank notes, “less than 20 percent of low- and middle-income countries have modern data infrastructure [...]. Even where nascent data systems and governance frameworks exist, a lack of institutions with the requisite administrative capacity,

decision-making autonomy, and financial resources holds back their effective implementation and enforcement” (World Bank, 2021, 13) (Figure 4).

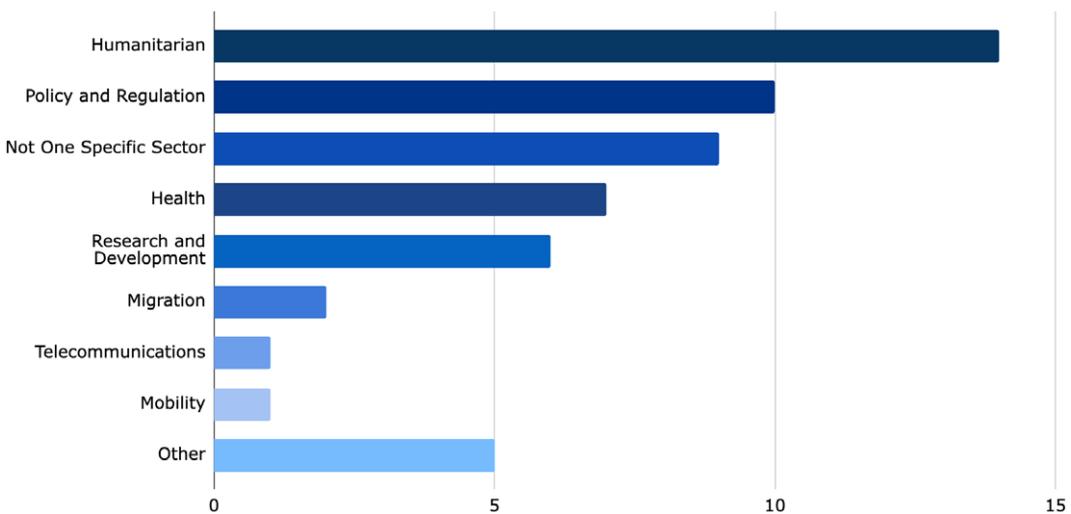
*Sectoral diversity:* As mentioned previously, this research was initially conducted to inform the UN HLCF Group, which is primarily focused on issues related to the UN’s programs and strategies for sustainable development. Given the group’s focus on humanitarian and development issues, the research prioritized frameworks and policies that were most relevant to those areas (Figure 5).

**3.2. U.N. agencies sub-sample**

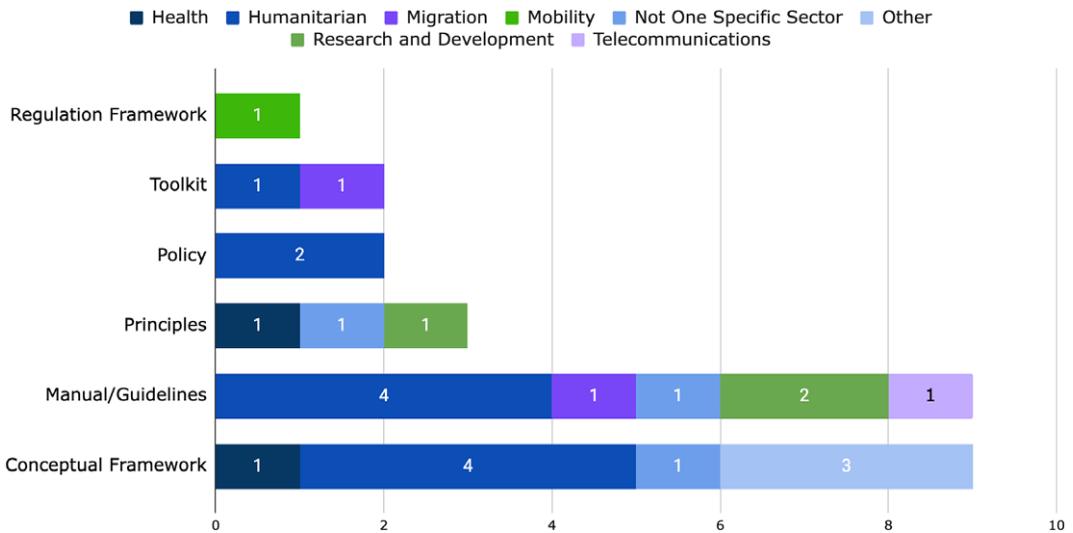
Given the initial intent to inform U.N. decisions on global data governance, it is worth declaring the weight the UN-system agencies represent in the sample. Of the frameworks analyzed, 26 correspond to UN-system data governance frameworks. Twenty-five of those have a global scope except for the Pan



*Figure 4. Geographical scope.*



*Figure 5. Sectoral diversity.*



**Figure 6.** Types and sectors of UN-system data governance frameworks.

American Health Organization's (PAHO) National Data Governance Framework, since PAHO serves as the Regional Office for the Americas of the World Health Organization. Of the U.N. agencies sub-sample, 69% are conceptual frameworks or guidelines, of which 42% are in the humanitarian sector (Figure 6).

### 3.3. Analytical framework

Once the empirical sample was assembled, the frameworks were analyzed through a conceptual prism consisting of the following "Template of Analysis."

---

**Purpose:** Purpose serves as the guiding objective of data governance frameworks. It illustrates the reason why a framework is needed in the first place, identifying a gap, and indicates the value the framework wants to bring about by filling that gap.

*Purpose:* Does the governance framework clarify its goals and objectives?

---



---

**Principles:** Principles serve as the guidance for a governance framework, ensuring that all activities are aligned with specific commonly agreed criteria and allows for easier interpretation.

*Principles:* Are there principles to guide the framework, and what are they?

*Anchoring:* Is the legal basis or other anchor documentation upon which the policies and principles built sufficiently explained? What is the nature of that basis?

---



---

**Data:** Describing and defining the data handled by the organization facilitates the understanding of the framework and pushes the organization to determine and justify the data they access and use. It also helps identify the data they seek to govern within the data value chain/data lifecycle.

*Data description:* Do the frameworks define the data they oversee? Is it personally identifiable data or not?

*Data lifecycle:* Does the framework describe the value-chain of data and the benefits and risks at each stage? (e.g., data localization).

---

---

**Processes and practices:** Operationalize the framework and ensure the principles are supported, and the processes are undertaken and monitored.

*Governance roles:* Does the framework explain any roles and functions that are tasked with the implementation of the framework? (e.g., code of conduct, data sharing agreement).

*Tools:* What tools and practices are specified to implement the framework?

*Monitoring and evaluating:* What monitoring and evaluation mechanisms are implemented?

---

### 3.4. Limitations

Despite the comprehensive nature of this research, there are several limitations to consider.

Firstly, for feasibility, the study focused on data governance frameworks created within the last 10 years (2013–2022), potentially missing out on valuable insights from earlier frameworks that could inform current practices. However, the research team considers that analyzing recent frameworks would offer the most pertinent and current insights on data governance practices. While the exclusion of older frameworks may omit some valuable insights, this limitation is deemed acceptable given the study's goal to provide current and relevant information.

Secondly, while the research aims to provide a comprehensive representation of data governance frameworks, the geographical scope is limited in terms of representation from the Global South. This may impact the generalizability of the findings, particularly for low- and middle-income countries.

Thirdly, the research prioritized public sector efforts over private-sector-oriented ones, cognizant of how the public sector plays a critical role in setting policies and regulations for data governance. However, the importance of the private sector in this area cannot be overlooked, and further research may be necessary to fully understand data governance in the private sector context.

Finally, the research was initially conducted to inform the UN HLCP Group, which is primarily focused on issues related to the UN's programs and strategies for sustainable development. While the sample includes a wide range of frameworks and approaches, encompassing principles, guidelines, and regulatory frameworks and standards, the prioritization of frameworks and policies relevant to humanitarian and development issues may limit the applicability of the findings to other sectors.

In conclusion, the study remains relevant to a wide range of stakeholders interested in data governance. Indeed, while the study's focus and limited scope may restrict the generalizability of the findings, the research team implemented a systematic and rigorous approach to provide valuable insights into contemporary practices and challenges in data governance. As such, the study's outcomes can inform and guide future research and policy-making endeavors in this critical field, demonstrating its significance and usefulness to various audiences.

## 4. Main Findings

The following section summarizes the key findings of the research with respect to existing data governance frameworks leveraging the previously described analytical framework.

### 4.1. Purpose

All of the reviewed data governance frameworks include an explanation of the overall purpose of the document. From the sample, there is great variety in the scope of the purposes identified by the different frameworks. This research identified two types of purposes: those that refer to specific cases and sectors where data are used, and those that aim to improve data governance in general. Often, the latter are pursued by national or local governments seeking to achieve responsible use and reuse of data.

As for the former, for instance, the International Committee of the Red Cross (ICRC)'s "Handbook on Data Protection in Humanitarian Action" seeks to raise awareness and assist humanitarian organizations

in ensuring that they comply with existing personal data protection standards in carrying out humanitarian activities specifically. As for the latter, on the other hand, the Personal Information Charter developed by the Foreign, Commonwealth & Development Office (FCDO) in the United Kingdom has very data-focused objectives that do not refer to one specific field (e.g., the humanitarian sector). Indeed, the charter provides the standards people can expect from the FCDO when they ask for, or hold, people's personal information.

Overall, an overarching purpose this research identified across all frameworks, regardless of their scope, is to balance the tension between the importance of *protection* against unauthorized collection and potential misuse of data versus the wider *promotion* of data for advancing various public interest goals. In fact, the frameworks identified often stemmed from the dual realization that data are indeed very valuable and potentially beneficial for public purposes, on the one hand, and that at the same time it poses a series of challenges and risks across its life cycle. Indeed, the promotion of data has the potential of bringing about a series of public benefits, spanning from more efficient mobility (Lau, 2020) to personalized healthcare (Morgan, 2021), from improved waste management (Abdallah et al., 2020) to more accessible education (Marchant, 2021). However, the misuse of data can result in potential issues including social exclusions (O'Neil, 2016) and injustices (Couldry and Mejias, 2019), wasted time and resources (Redman, 2016), as well as privacy (Cohen, 2012) and legal concerns (Rodrigues, 2020).

As a result, many governance frameworks seek to develop a variety of approaches aimed to leverage the new opportunities data presents, while avoiding the risks of its misuse. This appears to be a balancing act that is necessary but difficult to accomplish. Indeed, these efforts can result as being fragmented and difficult to operationalize.

One way to tackle such fragmentation is through data stewardship, which The GovLab defines as “policies, functions and competencies to enable access to and reuse of data for public benefit in a systematic, sustainable, and responsible way” (Verhulst, 2021a). Similarly, the Ada Lovelace Institute defines it as “responsible use, collection and management of data in a participatory and rights-preserving way” (Ada Lovelace Institute, 2021). The concept of stewardship has been used by Nobel Prize-winning economist Elinor Ostrom to describe the governance of common resources (Ostrom, 2012). When considering data as a common good, Ostrom's design principles and conceptualization of stewardship can be a useful tool to find a balance between data protection and data promotion. Data stewards can play a crucial role in steering the process of using data and the insights it can generate by deciding who has access to data, for what purposes and to whose benefit (Open Data Institute, 2022), ultimately addressing society's biggest questions and challenges in a systematic and responsible way (The GovLab, 2018).

## 4.2. Principles

### 4.2.1. Main themes

In examining the principles across the governance frameworks in the sample, this research identified three main themes:

First, trust is a main theme of the principles across the governance frameworks analyzed. Trust is essential both as a right in and of itself and moreover to enable the adoption and widespread usage of technologies and data platforms. Accordingly, a large number of the frameworks under examination emphasize trust as a central principle, both on ethical and practical grounds. An example can be found in the Privacy Impact Assessment developed by the Office of the Australian Information Commissioner, which emphasized a value proposition to strengthen community trust in the data initiatives carried out by the government. Another example, at the international level, are the WHO Data Principles. These highlight that their aim is to “provide a foundation for continually reaffirming trust in WHO's information and evidence on public health” (World Health Organization (WHO), 2020, 1).

Second, protecting citizen and user privacy rights emerges as one of the central principles in the frameworks analyzed in the sample, spanning across sectors and geographies. For example, the International Organization for Migration's Data Protection Manual emphasizes a core value proposition to “assist IOM staff to take reasonable and necessary precautions in order to preserve the confidentiality of

personal data and ensure that the rights and interests of IOM beneficiaries are adequately protected” (International Organization for Migration, 2015, 3).

The third central theme arising is the use of data for public interest. A large number of frameworks under examination emphasize the importance of increasing the scope of use of data so that it can be deployed more widely, in service of various public interests. For instance, the Data Sharing Policy of Médecins Sans Frontières (MSF) emphasizes that the organization’s repository of data “can potentially be of value to researchers working in public health” (Médecins Sans Frontières (MSF), 2013, 4).

4.2.2. Most used principles

This research identified the following three findings with regards to principles for data governance used in the frameworks examined in the sample: (a) lack of clarity and harmonization of meanings, (b) different meanings in different countries, and (c) association with Fair Information Practice Principles (Federal Privacy Council, n.d.) (Figure 7).

As for the former, this research identified a lack of clarity and harmonization of meanings among the principles analyzed. A large number of the frameworks include similar principles, but use different

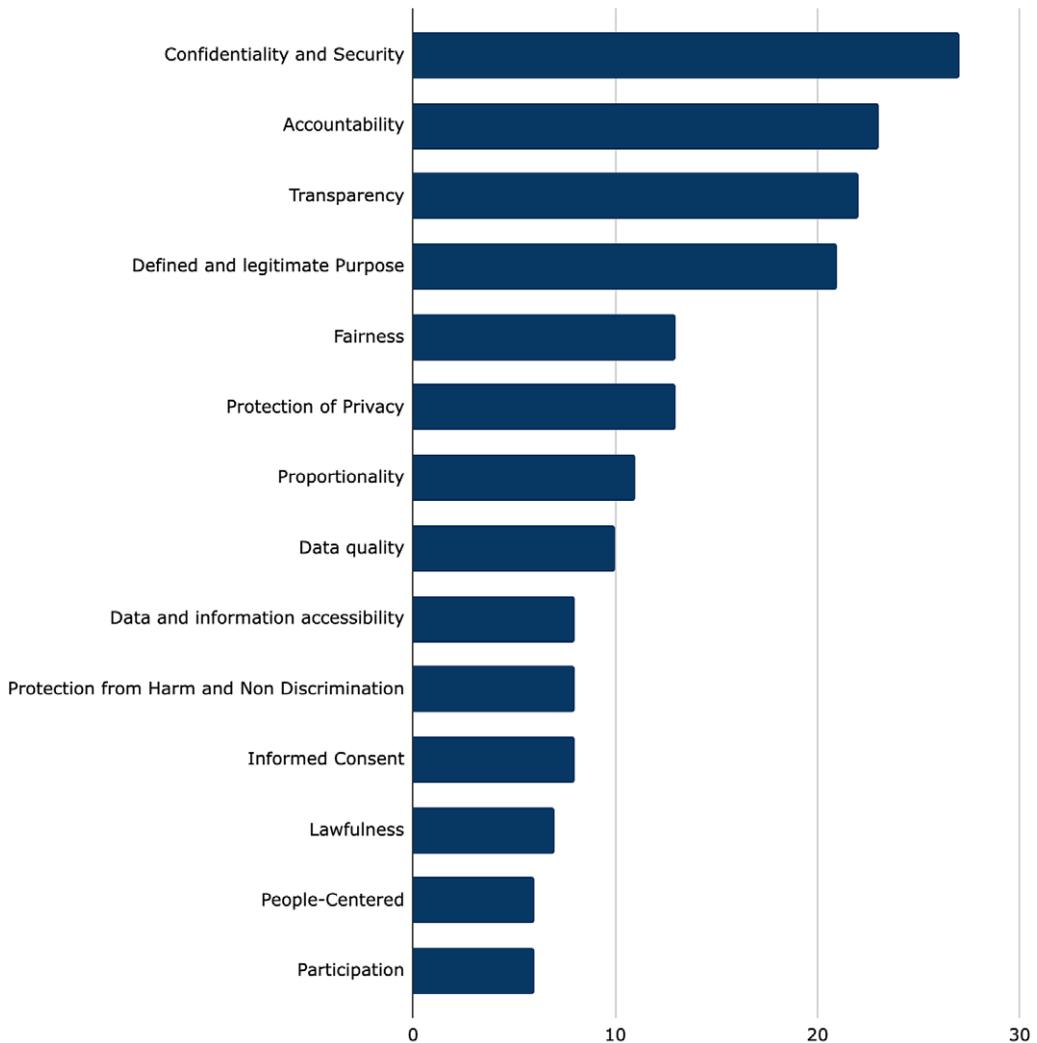


Figure 7. Most used principles.

Principles for Processes	Principles for Decisions	Principles for Data Handling
<ul style="list-style-type: none"> <li>• Transparency</li> <li>• Accountability</li> <li>• People-Centered</li> <li>• Fairness</li> <li>• Participation</li> <li>• Lawfulness</li> </ul>	<ul style="list-style-type: none"> <li>• Transparency</li> <li>• Proportionality</li> <li>• Defined Purpose</li> <li>• Accountability</li> <li>• People-Centered</li> <li>• Fairness</li> <li>• Protection from Harm</li> <li>• Non-Discrimination</li> <li>• Participation</li> </ul>	<ul style="list-style-type: none"> <li>• Confidentiality and Security</li> <li>• Proportionality</li> <li>• Data and Information Accessibility</li> <li>• Protection of Privacy</li> <li>• Lawfulness</li> <li>• Informed Consent</li> <li>• Data Quality</li> </ul>

*Figure 8. Categories of principles.*

nomenclature (thus increasing the challenges of cross-framework comparison and analysis). For instance, the principle of “equity” is absent in many frameworks, but it seems to be implied under the principles of “fairness” or “nondiscrimination.”

Moreover, a large number of frameworks do not separate principles, but group them together. Examples of principles often mentioned together are “necessity and proportionality” and “legitimate and fair processing.”

As for the second, this research found that principles can have different meanings in different countries. The difference in nomenclatures—and substantive definitions of similar concepts—is especially noteworthy across countries and geographic regions. For instance, “privacy” is defined and applied differently in different countries. How can that difference be reconciled under a global data governance framework?

Finally, many of the most used principles are associated with the Fair Information Practice Principles. Among those, this research found: confidentiality and security, purpose specification, transparency, accountability, data and information accessibility, data quality, proportionality, and participation.

#### 4.2.3. Processes, governance, and handling principles

This research identified different categories of principles that sought to inform and steer different aspects of the data governance life cycle that is processes, decisions, and data handling (albeit there is overlap): Principles for processes include principles whose aim is to shape the processes followed to arrive at certain governance decisions. Principles for decisions include those principles that aim to shape the governance decisions themselves. Principles for data handling aim to influence the way that data are processed and handled (Figure 8).

#### 4.3. Anchoring

Overall, 63% of the frameworks analyzed mention anchoring documents that they built on and referred to as a point of reference. Of those, 59% built on international human rights norms and principles. This means that only 39% of the analyzed approaches explicitly mention universal human rights frameworks. This research observed that anchor documents are mainly referred to as starting points instead of binding documents to comply with. In this sense, none of the analyzed frameworks mention roles or groups responsible for overseeing compliance with said anchor document. However, there is a relevant caveat in this context since many actors that authored the analyzed frameworks are international and supranational organizations that might possess specific privileges and autonomy to act and decide (Reinisch, 2009) and, therefore, have freedom in compliance with potential anchor documents.

Overall, the analysis identified three main types of frameworks in relation to anchor documents:

First, there are the frameworks that refer to international data and/or human rights protection standards as anchoring documents. An example of a framework that refers to international data protection standards is the “Data Strategy of the UN Secretary-General for Action by Everyone, Everywhere with Insight,

Impact and Integrity,” which emphasizes respect for human rights as well as international standards, such as the “UN Personal Data Protection and Privacy Principles” (United Nations, 2020, 60). Another example is the “Signal Program on Human Security and Technology” at the Harvard Humanitarian Initiative, which builds on rights identified in the Universal Declaration of Human Rights (UDHR), the International Covenants on Civil and Political Rights (ICCPR), and other instruments of humanitarian law rights that apply to all people “regardless of the use of any specific technology” (Harvard Humanitarian Initiative, 2015, 8).

Second, certain frameworks refer to previously established privacy legislation as anchoring documents. An example of a framework that builds on pre-existing privacy regulation is the Australian Information Commissioner’s Guide for Privacy Impact Assessment, which evaluates compliance to previously established privacy legislation such as the Privacy Act 1988. Another example is the United Kingdom’s Personal Information Charter, which seeks to ensure that all personal information is treated in accordance with the UK General Data Protection Regulation and the Data Protection Act from 2018.

Finally, a minority of the frameworks considered did not specifically refer to any legal basis. These were mainly collections of principles and general data governance guidelines, such as the “Data Privacy, Ethics and Protection: Guidance Note on Big Data for Achievement of the 2030 Agenda” developed by the United Nations Development Group (UNDG). This document sets out general guidance on data privacy, data protection and data ethics for the UNDG concerning the use of big data collected in real time by private sector entities, and was shared with UNDG members for the purposes of strengthening operational implementation of their programs to support the achievement of the 2030 Agenda. The guidelines do not refer to any legal basis, and provide a minimum base for self-regulation.

#### 4.4. Data description

This research conducted an analysis of the definition and treatment of data within various frameworks, focusing on data description and the data lifecycle. Regarding the classification of data, the findings are as follows:

Among the frameworks examined, 51% clearly specify the type of data they intend to govern. A significant majority of these frameworks primarily deal with the oversight of personal data, with approximately 90% of them focusing on personally identifiable information. While there is no unanimous consensus on a standardized data definition, several frameworks adopt a definition similar to that proposed by the UN World Food Programme, which states that “Personal data is any information relating to an individual that identifies the individual or can be used to identify them” (UN World Food Programme (WFP), 2016, 2). This definition aligns with the definitions used by other U.N. agencies such as UNICEF, UNFPA, and the International Organization for Migration.

On the other hand, 49% of the frameworks have an ambiguous or incomplete definition of data. These frameworks often provide a general definition of data but do not offer a specific definition tailored to the data they oversee. For example, the Responsible Data Management Training Pack by Oxfam adopts the definition formulated by the United Nations Economic and Social Commission for Asia and the Pacific (UNESCAP), which defines data as “the physical representation of information in a manner suitable for communication, interpretation, or processing by human beings or by automatic means. Data may be numerical, descriptive or visual” (Oxfam, 2015, 6).

In general, the analysis reveals that the governance frameworks lack recognition of emerging data definitions and types. Particularly, important concepts such as sensitive data and synthetic data are often absent from these frameworks. Furthermore, there seems to be a lack of exploration of relational data, which would necessitate a more comprehensive discussion on collective or community rights.

#### 4.5. Data lifecycle

When understanding data as a global asset, the concept of the “data lifecycle” is helpful for estimating the value of data. Value emerges in data transformation, from data collection, processing, and analysis into

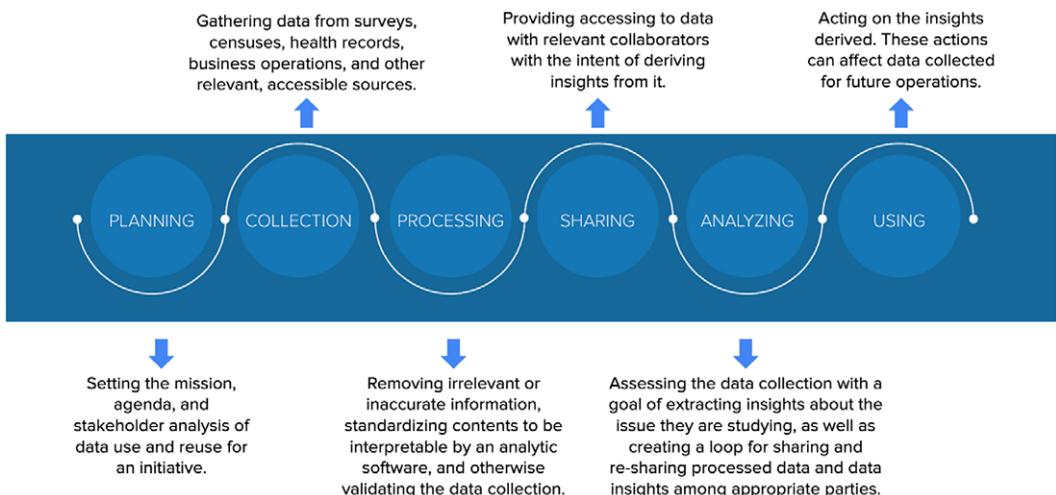
digital intelligence so that it can be monetized for commercial purposes or used for social objectives (UNCTAD, 2021, 17). This process identifies the particular characteristics and requirements when managing data in each stage. The decisions at every stage of the data life cycle will vary, depending on the type of data and their proximity to features of public goods. Therefore, addressing data governance through a data lifecycle approach is helpful because it enables a comprehensive analysis of how data should be overseen at various stages and create value from data use and reuse in a safe and equitable manner. Figure 9 presents a graphic representation of the data lifecycle.

Upon analyzing the selected frameworks from a data lifecycle perspective, it was observed that only 24% of the frameworks recognize the significance of the data lifecycle approach. These frameworks actively develop and provide specific recommendations for each stage of the data lifecycle while acknowledging the distinct requirements associated with each stage. A notable example is the ASEAN Data Management Framework, which was formulated to support the Data Life Cycle & Ecosystem, one of the strategic priorities outlined in the ASEAN Framework on Digital Data Governance. This framework delineates data governance principles across various stages of the data lifecycle, such as data collection, use, access, and storage. It also offers appropriate recommendations for safeguarding different types of data within organizations throughout the data lifecycle.

Similarly, the International Civil Aviation Organization (ICAO) has incorporated the data lifecycle concept into its Doc 8126 Aeronautical Information Services Manual from a sector-specific perspective. This manual acknowledges distinct stages within the Aeronautical Information Management concept, encompassing data acquisition, processing (including validation, verification, and management), provision of access to aeronautical information through information services, and consumption of aeronautical information by end users. The framework provides specific recommendations tailored to each stage.

Another illustration is the Responsible Data for Children Synthesis conceptual framework developed jointly by UNICEF and The GovLab. This framework delineates the data lifecycle across six broad stages: planning, collecting, storing and preparing, sharing, analyzing, and using. It offers actionable insights to facilitate responsible practices when working with children's data throughout the data lifecycle.

Among the analyzed frameworks, an additional 38% mention the data lifecycle approach either directly or implicitly but provide only partial or general recommendations. In some instances, the coverage of the data lifecycle stages is limited. Typically, existing governance frameworks primarily focus on data usage for the intended purpose during the planning stage. Conversely, only four of the



**Figure 9.** Data lifecycle by The GovLab.

analyzed frameworks address data localization requirements during the processing stage. There is also a notable absence of frameworks that address the data reuse stage comprehensively.

For example, the Data Security Law of the People's Republic of China recognizes the data lifecycle by including data handling activities such as collection, storage, use, processing, transmission, provision, and disclosure. However, this law does not provide specific recommendations or guidelines for each stage of the data lifecycle.

Similarly, the General Data Protection Regulation (GDPR) does not explicitly differentiate between the stages of data processing, as it defines processing as any operation performed on personal data, including collection, recording, organization, storage, adaptation, retrieval, consultation, use, disclosure, alignment, combination, restriction, erasure, or destruction. Although the GDPR covers various phases of the data lifecycle within the processing definition, it does not offer stage-specific recommendations. The exception to this is the chapter on "transfers of personal data to third countries or international organizations," where the focus is on data sharing.

Among the frameworks with limited scope, the Data Sharing Policy of Médecins Sans Frontières, a NGO, stands out. While this policy acknowledges the different stages of the data lifecycle, its sole concern is the data-sharing stage.

Lastly, 38% of the frameworks neither explicitly identify nor acknowledge the data lifecycle and its distinct stages as a basis for their approach to data governance.

#### 4.6. Processes

A well-functioning data governance framework should define the roles, responsibilities, and associated compliance procedures to safely share, use, and reuse data by all stakeholders. This research identified three types of processes: (a) process to develop the governance framework (b) process to identify and create new professions and functions, and (c) processes to monitor and evaluate.

Regarding the process to develop the governance framework, this study has observed notable advancements toward more participatory approaches, especially from a multisectoral perspective. Among the sampled frameworks, approximately 18% demonstrated the utilization of multi-stakeholder and multisectoral approaches in the formulation of their frameworks, particularly prevalent within intergovernmental institutions and international coalitions. For instance, the OECD's Recommendation of the Council on Health Data Governance was the outcome of a collaborative effort involving the Committee on Digital Economy Policy, the Health Committee, the former Working Party on Security and Privacy in the Digital Economy (now the Working Party on Data Governance and Privacy), and the former Health Care Quality Indicators Expert Group (OECD, 2016).

Nevertheless, it is worth noting that data subjects were often not included in these participatory processes. In fact, it is worth noting that only 5% of the frameworks followed a participatory process that included data subjects when defining the data governance approach. Among these, particularly illustrative example of greater participation and inclusiveness of data subjects is the Data Governance Framework for New Zealand, wherein the government engaged in co-designing the framework with the Māori community (data subjects) to ensure the incorporation of Māori needs and interests in data governance. Similarly, UNESCO's Internet Universality Indicators underwent a comprehensive development process involving multiple phases, rounds of consultations, and consultative meetings and workshops at international, regional, and national events, actively engaging a diverse array of stakeholders, including civil society and individuals.

However, in general, the remaining frameworks displayed shortcomings with regard to inclusivity, highlighting a need for improvement in this aspect.

Moving on to the process of identifying and creating new professions and functions, it seems that there is a pressing need for the establishment of specific roles and binding responsibilities within data governance frameworks. Approximately, 22.4% of the frameworks have implemented dedicated governance bodies with distinct functions to oversee the implementation and enforcement of the frameworks. These bodies assumed various titles, including Data Protection Officers and Chief Data Officers. For

instance, the Guidance on the Protection of Personal Data of Persons of Concern to United Nations High Commissioner for Refugees (UNHCR) introduced novel roles and responsibilities such as the Data Controller, Data Protection Officer, Inspector General's Office, and Ethics Office, each tasked with specific duties to ensure compliance with data protection policies (United Nations High Commissioner for Refugees [UNHCR], 2018). The GDPR, as another example, mandates that each member state appoint an independent public authority responsible for monitoring compliance with the regulation. Additionally, a European Data Protection Board, comprising heads of supervisory authorities from each member state, was established. Furthermore, the California Consumer Privacy Act (CCPA) established the California Privacy Protection Agency, empowered with comprehensive administrative authority to implement and enforce the CCPA, alongside the specification of a five-member board with expertise in privacy, technology, and consumer rights (CCPA, 2018).

However, it is worth highlighting that a considerable proportion of frameworks (24.1%) simply appended additional data governance functions to existing agencies or authorities, often without adequate capacity-building and expertise development. Notably, frameworks such as the Office for the Coordination of Humanitarian Affairs (UN OCHA) data responsibility guidelines and Oxfam's Responsible Program Data Policy assigned the responsibility of overseeing framework implementation to existing entities, including OCHA Centre for Humanitarian Data and Oxfam Country Directors, respectively (OCHA, 2021; Oxfam, n.d.). Similarly, UNICEF's Procedures for Ethical Standards in Research, Evaluation, Data Collection and Analysis by UNICEF states that Country Representatives, Regional Directors, and Heads of Divisions must ensure and maintain the highest ethical standards in all evidence generation endeavors (UNICEF Division of Data, Research and Policy, 2015, 2) by implementing the procedures laid out in the framework.

A small number of other frameworks (19%) recommend the establishment of new roles and governance bodies, but they often do not actually include provisions for creating them as part of nonbinding recommendations. For instance, the OECD's Recommendation of the Council on Health Data Governance provides vague instructions to existing governance bodies given its supranational nature. It recommends governments engage with relevant experts and organizations to develop mechanisms to implement the framework. The document also encourages NGOs to follow the recommendation when processing personal health data for health-related purposes that serve the public interest. Another example is the Association of Southeast Asian Nations (ASEAN) Data Management Framework (DMF) advises members to define roles and responsibilities for implementing each action described in the framework: "To develop and implement the 6 foundational components of the DMF, an organization is required to identify and determine different roles and responsibilities in order to ensure adoption, operation, and compliance, in accordance with business needs" (Association of Southeast Asian Nations (ASEAN), 2021, 12). However, in both instances, the recommendations are not legally binding for member states or non-governmental institutions.

From the sample, another category of frameworks (34.5%) emerges wherein information about governance mechanisms and the proposition of new roles or tasks are entirely absent. However, this absence can primarily be attributed to the nature and scope of the documents in question. Typically observed in principles-based frameworks, such as the Gemini Principles, the CARE Principles for Indigenous Data Governance, or the UN High-Level Committee on Management Personal Data Protection and Privacy Principles, these frameworks prioritize articulating values and principles for data management rather than explicitly outlining governance roles. For instance, the Risks, Harms and Benefits Assessment by UN Global Pulse is primarily focused on serving as a compliance mechanism for data privacy and protection, which limits its role to providing governance recommendations within its defined scope. As a result, these frameworks do not specify or suggest the establishment of governance roles, as their primary emphasis lies on guiding principles and values.

Lastly, with regard to processes related to monitoring and evaluation, 41.4% of the frameworks failed to establish or mention the need for monitoring mechanisms, which are vital components for the effective operationalization of any framework. For example, the "Handbook on Data Protection in Humanitarian Action" acknowledges that international organizations possess complete autonomy in data processing

and compliance monitoring, essentially constituting their own jurisdiction (International Committee of the Red Cross [ICRC], 2021).

A smaller portion of frameworks (29.3%) recommend the establishment of monitoring mechanisms, yet often lack clarity regarding the responsibility for their implementation. For instance, the International Organization for Migration's (IOM) Data Protection Manual, for instance, underscores the significance of oversight and compliance. However, the manual only advises the creation, without explicit appointment, of "an independent body to oversee the implementation of these principles and to investigate any complaints, and designated data protection focal points should assist with monitoring and training" (International Organization for Migration, 2015, 12). Similarly, the International Civil Aviation Organization's (ICAO) Aeronautical Information Services Manual suggests the establishment of monitoring and evaluation mechanisms but falls short in providing concrete instruments or mandates for their execution (ICAO, 2021): "States must implement well-documented surveillance processes by defining and planning inspections, audits, and monitoring activities on a continuous basis" (International Civil Aviation Organization, 2021, Section 2.8.1).

Conversely, the remaining 29.3% of the analyzed frameworks explicitly outline supervisory authorities responsible for overseeing and monitoring compliance. Some of these frameworks may correspond to regulatory frameworks or laws. For instance, Canada's Personal Information Protection and Electronic Documents Act (PIPEDA), for instance, establishes that the Office of the Privacy Commissioner of Canada will oversee compliance with the PIPEDA legislation. Further, the Commissioner may audit an organization's personal information management practices if the Commissioner believes that the organization has not followed a recommendation set out in the Act; it may ask for additional resources to monitor implementation. Similarly, the GDPR establishes in Articles 41 and 42 that a supervisory authority for monitoring compliance may be carried out by a body with an appropriate level of expertise in relation to the subject matter of the code and is accredited for that purpose. That way, the Member States, the supervisory authorities, the Board, and the Commission shall encourage compliance with this regulation of processing operations by data controllers and processors (General Data Protection Regulation, 2016). Another noteworthy example is the recommendation of the OECD council on health data governance. In this case, the framework provides detailed guidelines to implement monitoring and evaluation mechanisms, such as assessing whether the uses of personal health data have met the intended health-related public interest purposes and brought the benefits expected by (a) pursuing a periodic review of developments in personal health data availability, the needs of health research and related activities, and public policy needs; and (b) following a systematic assessment and updating of policies and practices to manage privacy, protection of personal health data and security risks relating to personal health data governance.

In addition, it is worth noting that none of the frameworks under study established evaluation mechanisms, which makes it difficult to assess their impact or effectiveness.

#### **4.7. Practices**

This benchmarking exercise examined the sample frameworks for their practices, which aimed to analyze their practical approach, and how they translate the theory, values, and principles into practice. From applied templates and checklists to the description of processes, sometimes imprecise, four trends were observed:

First, there is a plethora of practices but they often remain vague in terms of implementation. Indeed, 38% of the frameworks analyzed recommend and describe, often in a detailed fashion, good practices for data governance. However, in most cases, the frameworks do not include specific recommendations or guidelines for who, how, and when these practices should be implemented. For instance, The California Consumer Privacy Act (CCPA), for example, provides guidance to businesses on how to inform consumers of their rights under the CCPA, how to handle consumer requests, how to verify the identity of consumers making requests, and how to apply the law as it relates to minors. It stipulates the processes businesses need to implement to follow the CCPA while making it easier for consumers to exercise their

CCPA rights (California Consumer Privacy Act, 2018). Similarly, the UNICEF and UNFPA Policy on Personal Data Protection. The policy describes good practice when collecting data from individuals and explains how to notify the data subject. It presents the information that shall be provided to each identified data subject within a reasonable period when personal data are collected by UNICEF or UNFPA (as controller), taking into account the logistical constraints both organizations face (UNICEF, 2020). Another example is the Privacy Impact Assessment Toolkit from the Office of the Australian Information Commissioner describes a detailed 10-step process for undertaking a privacy impact assessment to apply the toolkit. Likewise, the data responsibility guidelines of the UN Office for the Coordination of Humanitarian Affairs (UN OCHA) recommend eight actions to be implemented at the system-wide, sector, and organizational levels. The framework guides OCHA Staff to implement these processes at different levels. However, there is no clarity regarding the timing and the ways in which the actions should be implemented.

Second, there seems to be a need for developing practical tools, such as templates for sharing agreements. Data collection and data sharing are key components of the data life cycle, but rife with potential minefields (e.g., potential regulatory or privacy violations). Organizations, especially if they are under-resourced technically or financially, can greatly benefit from detailed guidelines such as templates and checklists to help guide their decisions. In the sample, only 29% of frameworks provided tools such as templates, checklists, or assessments.

One example is the International Organization for Migration's Data Protection Manual that provides a number of practical tools: templates and checklists. Templates of model consent forms, general contractual clauses to be inserted into contracts, and request forms for data subjects seeking access to their personal data; and checklists for data quality, data security, and data protection. Moreover, the USAID's Considerations for Using Data Responsibly offers applied tools for how to help guide discussions or navigate areas of responsible data practice that may be unclear. The tools range from the key events planning table to the benefits risk assessment, a worksheet to track and protect copies of sensitive data and IT security highlights checklist.

In the case of the United Nations Development Programme's (UNDP) Data Principles, although they primarily comprise guiding principles, they also provide practical resources to support the implementation of their recommendations. For instance, resources include an explanation of informed consent for safeguarding personal data, the Responsible Development Data Book for managing data responsibly, and the Mozilla Science Data Reuse Checklist for planning reusability and interoperability. Another example providing practical tools is the WFP Guide to Personal Data Protection and Privacy, which presents a Self-Assessment Compliance Checklist enabling personnel to gauge compliance with each element of the guidelines. Additionally, the guide offers Model Consent Forms that can serve as a basis for developing localized templates for obtaining informed consent and responding to beneficiaries' requests for accessing their data.

Third, noteworthy examples of best practices emerge in the realm of risk assessments and compliance mechanisms. Among the processes and tools delineated in the analyzed frameworks, a subset of risk assessments stands out as exemplifying responsible data management practices.

For example, the Risks, Harms and Benefits Assessment of the UN Global Pulse develops a two-steps assessment: the first one is a checklist, which is used as an initial assessment tool that identifies potential risks and helps evaluate whether a more comprehensive review should be conducted; and the second one consists of a detailed measurement of the likelihood, magnitude, and significance of impacts of a data innovation project if a medium or high risk was identified in the initial assessment. Furthermore, the Handbook on Data Protection in Humanitarian Action by the Brussels Privacy Hub (VUB) and ICRC offers a Data Protection Impact Assessment (DPIA) tool to identify, evaluate and address the risks to personal data arising from a project, policy, program, or another initiative. It includes a step-by-step guide for humanitarian organizations to conduct it. It also has a template for a DPIA report. Likewise, within the suite of tools provided in the Considerations for Using Data Responsibly at USAID, a benefits risk assessment tool is developed. This tool serves the purpose of assessing the potential benefits and risks associated with data collection, use, and sharing. A key aspect of conducting a thorough evaluation of

risks and benefits lies in ensuring the involvement of relevant stakeholders, including those from whom data are collected.

Finally, this research found predominantly nonbinding frameworks. With few exceptions (e.g., data protection regulatory frameworks and laws), most recommendations are not binding. This may limit their effectiveness and the extent to which they are operationalized. Their list of recommendations and good practices are “suggestions” or “proposals.” Furthermore, 34% of the data frameworks reviewed do not provide detailed processes or explicit practical tools.

#### **4.8. Key takeaways**

This section presents the key takeaways from the research article’s analysis of the data governance ecosystem within sectors closely related to international development. The findings provide insights into six key elements: purpose, principles, anchoring, data description and lifecycle, processes, and practices (see Section 3.3). The analysis revealed a series of patterns within each element, shedding light on the current state of data governance frameworks. These findings lay the groundwork for the final recommendations illustrated in the following and final section.

##### *4.8.1. Purpose*

The findings reveal that every data governance framework reviewed in the research article explicitly stated its purpose. The frameworks analyzed aimed to establish a clear direction and guide decision-making processes regarding data governance. The goals and objectives outlined in these frameworks varied, with some focusing on specific sectors and others emphasizing broader aims. However, a notable emerging purpose identified across all frameworks was the reconciliation of the tension between data protection and the increasing use of data for societal goals. This reflects the ongoing challenge of striking a balance between safeguarding individual privacy and leveraging data for societal benefits.

##### *4.8.2. Principles*

The research article identified three main themes in relation to the principles mentioned in the reviewed data governance frameworks. Firstly, the importance of trust was emphasized, highlighting the need to foster trust between data holders, data users, and individuals whose data are being accessed. Secondly, individual rights and interests were recognized as crucial, acknowledging the significance of protecting privacy and ensuring data subjects have control over their personal information. Lastly, the principle of public interest was highlighted, emphasizing the need to consider broader societal well-being when making decisions regarding data governance.

Among the principles used in the frameworks, confidentiality and security, accountability, and transparency were the most frequently mentioned. These principles build upon the well-established Fair Information Practice Principles (FIPPs) and provide a foundation for data governance frameworks. However, there seems to be a lack of clarity and harmonization of meanings, particularly across different countries. This lack of consistency makes it challenging to effectively implement and operationalize the principles. The research also identified three main types of principles: principles for processes, principles for decisions, and principles for handling data, each serving a unique role in guiding data governance practices.

##### *4.8.3. Anchoring*

In terms of anchoring, the research revealed that 63% of the analyzed frameworks mentioned specific documents that served as reference points or starting points for their development. These anchor documents were crucial for establishing a foundation and providing guidance in the formulation of the frameworks. However, it is important to note that these documents were often referred to as starting points rather than binding documents that frameworks must strictly comply with.

Among the various anchoring documents mentioned, 39% of the frameworks explicitly referred to universal human rights frameworks, indicating a recognition of the significance of human rights in the context of data governance. Additionally, two specific types of anchoring documents were identified in the sample: international data and/or human rights protection standards and previously established privacy legislation. However, it was also noted that a minority of frameworks did not explicitly refer to any legal basis, potentially posing challenges in terms of legal validity and compliance.

#### 4.8.4. Data description

Regarding data description, the research found that 51% of the analyzed frameworks provided clear definitions of the type of data they aimed to oversee. These frameworks recognized the importance of precisely defining the scope and nature of the data under consideration. However, 49% of the frameworks had an unclear or only partial definition of data. While these frameworks addressed data in general terms, they lacked specific definitions for the data they were designed to govern.

The research also highlighted the absence of emerging concepts such as sensitive and synthetic data, as well as concepts related to relational data, group privacy, and collective rights. These gaps indicate a need for further development and consideration of evolving data types and associated privacy concerns within the frameworks.

#### 4.8.5. Data lifecycle

In terms of the data lifecycle, the research findings indicate that 24% of the analyzed frameworks recognized and acknowledged the relevance of adopting a data lifecycle approach. These frameworks provided specific recommendations for each stage of the data lifecycle while acknowledging the unique needs and considerations associated with each phase. However, 38% of the frameworks only offered partial or general recommendations, often with limited scope or specific coverage of certain stages of the data.

A notable observation was the lack of frameworks that adequately addressed the data reuse stage, indicating a gap in understanding and guidance in this crucial aspect of data governance. Further attention and focus are needed to ensure comprehensive coverage of all stages of the data lifecycle within data governance frameworks.

#### 4.8.6. Processes

The research revealed that only 5% of the frameworks followed a participatory process that included data subjects in defining the data governance approach. This low percentage suggests a limited emphasis on involving those whose data are being governed in the decision-making process. In contrast, 22.4% of the frameworks established specific governance bodies responsible for overseeing the implementation of the frameworks. The remaining frameworks either added additional functions to existing agencies or authorities or provided vague high-level recommendations.

Furthermore, the research found that only 29% of the reviewed frameworks explicitly stated how supervisory authorities would monitor and evaluate compliance with the framework. Another 29% vaguely recommended the establishment of a monitoring mechanism, while 41% did not address or mention the need for monitoring mechanisms at all. This indicates a lack of clarity and consistency regarding the monitoring and evaluation of data governance frameworks, which may impact their effectiveness in practice.

#### 4.8.7. Practices

Regarding practices, the research identified that only 29% of the frameworks provided practical tools such as templates, checklists, or assessments. These frameworks recognized the importance of offering tangible resources to facilitate the implementation of data governance practices. However, 38% of the analyzed frameworks recommended and described good practices for data governance in a detailed

manner, without clearly defining who should implement these practices, how they should be implemented, and when they should be applied.

It was noted that, with a few exceptions such as data protection regulatory frameworks and laws, most of the recommended practices within the frameworks were not binding. This lack of binding nature may limit their effectiveness and the extent to which they are operationalized in practice. Greater clarity and specificity regarding the implementation of recommended practices would enhance their usability and impact within the data governance ecosystem.

## 5. Final Recommendations

Data governance embraces a wide range of elements and concepts without a unified or unique definition, which might create asymmetries when establishing a data governance framework. The analysis has shown a variety of purposes, approaches, and scopes at the local, national, and international levels. Davis proposes a definition aiming to gather multiple components of it: “Data governance concerns the rules, processes and behaviors related to the collection, management, analysis, use, sharing and disposal of data – personal and/or non-personal. Good data governance should promote benefits and minimize harms at each stage of relevant data cycle” (Davis, 2022, 12).

Well-designed data governance, according to the World Bank, can be defined as the framework that allows capturing the central values and purposes of an entity (country, international body, region, etc.) to leverage the synergies with multiple stakeholders while creating trust and promoting the use of data (World Bank, 2021, 10). Based on those definitions and building upon the main takeaways from the detailed analysis of 58 data governance frameworks, there is an opportunity for researchers, decision-makers, and other stakeholders to identify critical elements and follow good practices. Some of these may include:

*Considering data stewardship as a way to reconcile the tension between data protection and data promotion:* Moving forward, there may be a need to adopt a broader framework and concept of data stewardship. This would indeed allow to achieve and maintain the dual goal of protecting and promoting data in a more systematic, sustainable, and responsible way. Indeed, as mentioned in Section 4.1, data stewardship aims to make the use of data more responsible, systematic, and sustainable (Verhulst, 2021a); achieve the responsible and accountable use of common resources (Ada Lovelace Institute, 2021) allowing to make full use of data’s benefits and avoiding the social and economic harms that can stem from its misuse (Open Data Institute, 2022). In particular, it may be useful to (a) provide a legal, shared definition for global data stewardship, (b) rationalize and coordinate existing support to international data stewardship efforts, and (c) commission research and trials to assess the potential of a global data stewards association.

*Focusing on responsible reuse to unlock the socioeconomic value of data:* In recent years, the open data movement to improve public governance has grown significantly (The GovLab, 2016). As a consequence, increasing amounts of both public and private data have been made available to external stakeholders. However, although the frameworks analyzed in this research did aim to develop different ways to govern the use of data, they overall lacked a focus on the reuse of data—that is, the sharing of data across different domains. It may in fact be beneficial to integrate the concept of reuse in the development of a global data governance framework, so as to create shared approaches and standards with respect to the sharing of data amongst different stakeholders. In particular, it may be useful to (a) develop methodologies to define and measure the value of data, (b) develop structures to incentivize the “co-creation of value” (Mazzucato, 2019), (c) encourage data collaboratives,<sup>1</sup> and (d) identify and nurture data stewards, as further specified in Reflection 1 (Verhulst, 2020). Finally, in order for the reuse of data to be deemed

<sup>1</sup> In this research, “data collaboratives” are defined as “... an emerging form of public–private partnership that enables sharing and co-creation of value. They may involve, for instance, informal and time-bound collaborations between a company and an academic research group or civil society organization, and allow data to be re-purposed, typically in an anonymized form and with specific intent” (Verhulst, 2020).

responsible and consequently legitimate, it is crucial to create avenues for public assemblies and value the importance of social licenses (see Recommendation 7; Verhulst, 2021b).

*Harmonizing meanings to operationalize principles:* This research showed that there is an overall lack of clarity and harmonization of meanings across different countries, sectors, and organizations. This makes it difficult to operationalize the principles in a harmonized manner and at a global level. Whereas different contexts are bound to value different principles to some degree, there seem to be an overarching agreement on a series of data governance principles (see Section 4.2.2). These, however, seem to be defined differently by different organizations. It may be worth universalizing the principles to be embedded in a global data governance framework, so as to systematically operationalize them and ensure compliance across different regions and nations.

*Using broader anchoring frameworks to provide common North Stars:* As mentioned in Section 4.2, of all the approaches analyzed, only 39% explicitly mention global human rights frameworks. Moreover, anchor documents are mainly starting points, instead of binding documents to comply with. Finally, because of the loose nature of the “anchoring process,” none of the frameworks clearly mention responsible roles or identify processes for overseeing compliance with an anchor document. First, it seems that having broader frameworks—and not only, for instance, privacy-focused legal bases—related to universal human rights may be beneficial in developing a global data governance framework (MacFeely et al., 2022). Second, it seems important to establish clear levels of compliance required with such documents. Finally, based on those levels, it may be beneficial to answer the question of who oversees compliance, so as to ultimately materialize the relationship between the framework and its anchor document.

*Unifying key definitions of data and incorporate emerging concepts such as synthetic data:* Whereas it is important to keep in mind that a fixed definition of data may be more harmful than beneficial, mainly due to the ever-changing nature of both data and the technologies it relates to, it is also crucial to develop a series of mechanisms that allow flexibility and clarity of what people and organizations mean by “data.” In this sense, it may prove useful to incorporate emerging, flexible concepts such as synthetic data, as well as relational data, thick data, and sensitive data (with the latter having been increasingly more adopted lately). This indeed could enable the definitions to be more precise, without referring to the broad, general concept of data, and at the same time it may result in a malleable approach that could allow for the various data-related evolutions and developments to be assimilated.

*Adopting the data lifecycle approach to promote benefits and minimize harms:* Given the breadth of contexts in which data governance must be applied, it is beneficial to use a standardized framing to structure the needs, risks, and opportunities when handling data. As mentioned before, data governance results from multiple processes all aligned toward data promotion and protection. These processes can be hard to understand when viewed together, and although the data lifecycle is not linear, it could help to inform responsible data handling approaches better while promoting better and more impactful data management (The GovLab, 2021).

*Incorporating more participatory processes and collective agency to develop a data governance framework:* Participatory data governance occurs when organizations allow different constituents to contribute to the discussion and are accountable for their decisions to the public. To encourage transparency and accountability in data governance efforts, decision-makers should offer opportunities for scrutiny and input from data subjects. This policy feedback process is particularly relevant within the data governance discussion since it will allow obtaining the most value from data while protecting people from harm (The Digital Trade and Data Governance Hub, 2022). Therefore, public consultations on the design of policies and regulations could support transparency and stakeholder engagement (World Bank, 2021, 284) while fostering the social license of the process. The social license refers to the informal permissions granted to institutions such as governments or corporations by members of the public to carry out a particular set of activities (Shaw et al., 2020), in this case, the collection, sharing, and use of their data.

The vast majority of today’s participatory processes focus on protecting individuals’ rights. Yet these debates fail to consider the agency of data subjects as a collective. There are often massive asymmetries

between individuals and stronger stakeholders, such as the public or the private sector, that exploit their data while restricting its potential. To address these asymmetries, a new principle of digital self-determination is needed. Verhulst defines digital self-determination as “the principle of respecting, embedding, and enforcing people’s and people’s agency, rights, interests, preferences, and expectations throughout the digital data life cycle in a mutually beneficial manner for all parties involved” (Verhulst, 2023, 6). In the context of data governance, building symmetric relationships can help data subjects leverage their self-determination more effectively to exert greater control over how their data are used and reused (Verhulst, 2023). This is particularly relevant for underrepresented groups, which possess even lower bargaining power than other collectives.

*Investing in and creating new professions with specific roles and responsibilities:* Attracting data talent and promoting data stewardship are, key to ensuring compliance with data governance frameworks and fostering a culture around data collaboration and protection within organizations. The Third Wave of Data by The GovLab proposes a focus on new institutional arrangements to achieve a data-driven culture with particular attention on the role of the data steward—accountable data leaders that seek new ways to create value through cross-sector data collaboration (The GovLab, 2021). The analysis reveals the importance of having trained and dedicated individuals (whether chief data, chief privacy, or chief security officers, or the equivalent body) with specific functions and binding responsibilities for long-lasting, sustainable, and informed data actions.

*Improving accountability and transparency by defining oversight and compliance mechanisms:* There is a need to explicitly define monitoring and evaluation mechanisms linked to defining roles and responsibilities. An organization’s accountability can be measured by how it monitors and assesses its internal policies to manage, protect, and secure data effectively (Centre for Information Policy and Leadership, 2011) or simply by complying with the data governance policy. To do so, it is recommended to establish not only the mechanism but also who, when, how often, and how it should be implemented.

*Translate values and recommendations into practical tools:* In collaboration with diverse policy-makers and stakeholders, identify the most valuable tools to facilitate and accelerate the implementation of a data governance policy. The analysis identified three practical tools that should be considered to help data stewardship: model consent forms, checklists for data quality, data security and data protection, and data risks assessment step-by-step guidelines. These tools may indeed prove useful to guide implementation, document progress, and monitor compliance.

**Supplementary material.** The supplementary material for this article can be found at <https://doi.org/10.1017/dap.2023.24>.

**Acknowledgments.** This work would not have been possible without the support and guidance of many individuals throughout the research process. We would like to express our sincere gratitude to the peer reviewers for their valuable contributions and insights in reviewing this report. In particular, we would like to thank Dr. Susan Aaronson, Research Professor at George Washington University and Director of the Digital Trade and Data Governance Hub, Jos Berens, Data Policy Officer at United Nations Office for the Coordination of Humanitarian Affairs (OCHA), and Malar Veerappan, Program Manager and Senior Data Scientist at the World Bank. Finally, we would like to thank the United Nations High-Level Committee on Programmes (HLCP) Core Group on New Global Public Goods for their support and guidance throughout this research.

**Funding statement.** The GovLab received funding from the World Health Organization to develop an initial mapping of data governance regimes to inform the deliberations of the United Nations High-Level Committee on Programmes (HLCP) Core Group on New Global Public Goods.

**Competing interest.** The authors declare no competing interests exist.

**Author contribution.** Writing—original draft: S.M., N.A.G., E.W.; Writing—review and editing: S.G.V.

**Data availability statement.** The authors intend to share data via a repository on acceptance.

## References

Abdallah M, Talib MA, Feroz S, Nasir Q, Abdalla H and Mahfood B (2020) Artificial intelligence applications in solid waste management: A systematic research review. *Waste Management* 109, 231–246. <https://doi.org/10.1016/j.wasman.2020.04.057>

- Ada Lovelace Institute** (2021) *Participatory Data Stewardship*. Ada Lovelace Institute. Available at <https://www.adalovelaceinstitute.org/report/participatory-data-stewardship/> (accessed 15 December 2022).
- Anthony BJ** (2022) Toward a collaborative governance model for distributed ledger technology adoption in organizations. *Environment Systems and Decisions* 42, 276–294. <https://doi.org/10.1007/s10669-022-09852-4>
- Aaronson SA** (2022) Building trust in digital trade will require a rethink of trade policy-making. *Oxford Review of Economic Policy* 39(1), 98–109. <https://doi.org/10.1093/oxrep/grac046>
- Association of Southeast Asian Nations (ASEAN)** (2021) ASEAN Data Management Framework. Data Governance and Protection Throughout the Data Lifecycle. Available at <https://asean.org/wp-content/uploads/2021/08/ASEAN-Data-Management-Framework.pdf> (accessed 15 December 2022).
- California Consumer Privacy Act** (2018) *California Civil Code Division 3 Part 4*. Sacramento, California State Legislature: California Consumer Privacy Act of 2018.
- Centre for Information Policy and Leadership** (2011) *Accountability: Data Governance for the Evolving Digital Marketplace*. Hunton & Williams LLP. Available at <http://web.archive.org/web/20130523011828/>; [http://www.huntonfiles.com/files/webupload/CIPL\\_Centre\\_Accountability\\_Data\\_Governance\\_Paper\\_2011.pdf](http://www.huntonfiles.com/files/webupload/CIPL_Centre_Accountability_Data_Governance_Paper_2011.pdf) (accessed 15 December 2022).
- Chen R** (2021) Mapping data governance legal frameworks around the world. Findings from the Global Data Regulation Diagnostic. Policy Research Working Paper 9615. Available at <https://openknowledge.worldbank.org/server/api/core/bitstreams/0a248046-b7c9-59eb-a2e5-2d39e3a0b6be/content> (accessed 15 December 2022).
- Cohen J** (2012) What privacy is for. *Harvard Law Review* 126. <https://ssrn.com/abstract=2175406> (accessed 15 December 2022).
- Coyte D** (2020) *Common Governance of Data: Appropriate Models for Collective and Individual Rights*. Ada Lovelace Institute. Available at <https://www.adalovelaceinstitute.org/blog/common-governance-of-data/> (accessed 15 December 2022).
- Coyle D, Diepeveen S, Wdowin J, Kay L and Tennison J** (2020) *The Value of Data: Policy Implications*. Bennett Institute for Public Policy. Available at [https://www.bennettinstitute.cam.ac.uk/wp-content/uploads/2020/12/Value\\_of\\_data\\_Policy\\_Implications\\_Report\\_26\\_Feb\\_ok4noWn.pdf](https://www.bennettinstitute.cam.ac.uk/wp-content/uploads/2020/12/Value_of_data_Policy_Implications_Report_26_Feb_ok4noWn.pdf) (accessed 15 December 2022).
- Couldry N and Mejias U** (2019) *The Costs of Connection: How Data is Colonizing Human Life and Appropriating it for Capitalism*. Redwood City, CA: Stanford University Press.
- Davis T** (2022) Data governance and the datasphere. Literature review. *Datasphere*. Available at <https://www.thedatasphere.org/wp-content/uploads/2022/11/Data-governance-and-the-Datasphere-Literature-Review-2022.-Tim-Davies.pdf> (accessed 15 December 2022).
- Fay R** (2022) *Global Governance of Data and Digital Technologies: A Framework for Peaceful Cooperation*. Centre for International Governance Innovation. Available at <https://www.cigionline.org/articles/global-governance-of-data-and-digital-technologies-a-framework-for-peaceful-cooperation/> (accessed 15 December 2022).
- Federal Privacy Council** (n.d.) Fair Information Practice Principles (FIPPs). Available at <https://www.fpc.gov/resources/fipps/> (accessed 15 December 2022).
- General Data Protection Regulation** (2016) Available at <https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:32016R0679> (accessed 15 December 2022).
- Harvard Humanitarian Initiative** (2015) “The Signal Code,” A Human Rights Approach to Information during Crisis. Available at [https://hhi.harvard.edu/files/humanitarianinitiative/files/signalcode\\_final.pdf?m=1607469621](https://hhi.harvard.edu/files/humanitarianinitiative/files/signalcode_final.pdf?m=1607469621) (accessed 15 December 2022).
- International Civil Aviation Organization** (2021) *Aeronautical Information Services Manual, 7th Edn*. Available at [https://www.icao.int/NACC/Documents/eDOCS/AIM/8126\\_unedited\\_en%20Jul2021.pdf](https://www.icao.int/NACC/Documents/eDOCS/AIM/8126_unedited_en%20Jul2021.pdf) (accessed 15 December 2022).
- International Committee of the Red Cross [ICRC]** (2021) *Handbook on Data Protection in Humanitarian Action*. International Committee of the Red Cross. Available at <https://www.icrc.org/en/data-protection-humanitarian-action-handbook> (accessed 15 December 2022).
- International Organization for Migration** (2015) *IOM Data Protection Manual*. International Organization for Migration. Available at <https://publications.iom.int/books/iom-data-protection-manual> (accessed 15 December 2022).
- Lau J** (2020) Google Maps 101: How AI helps predict traffic and determine routes. The Keyword. Available at <https://blog.google/products/maps/google-maps-101-how-ai-helps-predict-traffic-and-determine-routes/> (accessed 15 December 2022).
- MacFeely S, Me A, Fu H, Veerappan M, Hereward M, Passarelli D and Schüür F** (2022) Towards an international data governance framework. *Statistical Journal of the IAOS* 38(4), 703–710. <https://doi.org/10.3233/SJI-220038>
- Marchant N** (2021) *Education technologies are making learning more accessible | World Economic Forum. The World Economic Forum*. Available at <https://www.weforum.org/agenda/2021/01/education-technology-accessibility-learning/> (accessed 15 December 2022).
- Mazzucato M** (2019) *Preventing digital feudalism by Mariana Mazzucato. Project Syndicate*. Available at <https://www.project-syndicate.org/commentary/platform-economy-digital-feudalism-by-mariana-mazzucato-2019-10> (accessed 15 December 2022).
- Médecins Sans Frontières (MSF)** (2013) MSF Data Sharing Policy. Available at [https://www.msf.org/sites/default/files/msf\\_data\\_sharing\\_policy\\_final\\_061213.pdf](https://www.msf.org/sites/default/files/msf_data_sharing_policy_final_061213.pdf) (accessed 15 December 2022).
- Medhora RP and Owen T** (2020) *A post-COVID-19 digital Bretton Woods by Rohinton P. Medhora & Taylor Owen. Project Syndicate*. Available at <https://www.project-syndicate.org/onpoint/digital-bretton-woods-new-global-governance-model-by-rohinton-p-medhora-and-taylor-owen-2020-04> (accessed 15 December 2022).
- Morgan B** (2021) *The future of healthcare personalization. Forbes*, May 3, 2021. Available at <https://www.forbes.com/sites/blakemorgan/2021/05/03/the-future-of-healthcare-personalization/> (accessed 15 December 2022).

- Mungai C, Robehmed S, Dahiya V, Arora C, Fusari M, Beltrán H, Pollio A, et al.** (2022) *Parables of AI in/from the Majority World*. Data & Society. Available at <https://datasociety.net/library/parables-of-ai-in-from-the-majority-world-an-anthology/> (accessed 15 December 2022).
- O’Neil C** (2016) *Weapons of Math Destruction: How Big Data Increases Inequality and Threatens Democracy*. New York: Crown.
- OCHA Centre for Humanitarian Data** (2021) OCHA Data Responsibility Guidelines. Available at [https://data.humdata.org/dataset/2048a947-5714-4220-905b-e662cbcd14c8/resource/60050608-0095-4c11-86cd-0a1fc5c29fd9/download/ocha-data-responsibility-guidelines\\_2021.pdf](https://data.humdata.org/dataset/2048a947-5714-4220-905b-e662cbcd14c8/resource/60050608-0095-4c11-86cd-0a1fc5c29fd9/download/ocha-data-responsibility-guidelines_2021.pdf) (accessed 15 December 2022).
- OECD** (2016) OECD Legal Instruments. Available at <https://legalinstruments.oecd.org/en/instruments/OECD-LEGAL-0433> (accessed 15 December 2022).
- Open Data Institute** (2022) *Bottom-Up Data Institutions: Mechanisms for Government Support*. Open Data Institute. Available at [https://theodi2022.wpengine.com/wp-content/uploads/2022/03/ODI-2022-Bottom-up-data-institutions\\_Mechanisms-for-government-support.pdf](https://theodi2022.wpengine.com/wp-content/uploads/2022/03/ODI-2022-Bottom-up-data-institutions_Mechanisms-for-government-support.pdf) (accessed 15 December 2022).
- Ostrom E** (2012) *Governing the Commons: The Evolution of Institutions for Collective Action*. Cambridge: Cambridge University Press.
- Oxfam** (2015) *Responsible Program Data Policy*. Oxfam International. Available at <https://policy-practice.oxfam.org/resources/oxfam-responsible-program-data-policy-575950/> (accessed 15 December 2022).
- Pisa M and Nwankwo U** (2021) *How Can Multilateral Organizations Strengthen Global Data Governance Practices?* Center for Global Development. Available at <https://www.cgdev.org/sites/default/files/how-can-multilaterals-strengthen-global-data-governance-brief.pdf> (accessed 15 December 2022).
- Redman TC** (2016) Bad data costs the U.S. \$3 trillion per year. *Harvard Business Review*. Available at <https://hbr.org/2016/09/bad-data-costs-the-u-s-3-trillion-per-year> (accessed 15 December 2022).
- Reinisch A** (2009) *Convention on the Privileges and Immunities of the United Nations*. United Nations – Office of Legal Affairs. <https://legal.un.org/avl/ha/cpiun-cpisa/cpiun-cpisa.html> (accessed 16 February 2023).
- Rodrigues R** (2020) Legal and human rights issues of AI: Gaps, challenges and vulnerabilities. *Journal of Responsible Technology* 4, 100005. <https://doi.org/10.1016/j.jrt.2020.100005>.
- Shaw JA, Sethi N and Cassel CK** (2020) Social license for the use of big data in the COVID-19 era. *NPJ Digital Medicine* 3, 128. <https://doi.org/10.1038/s41746-020-00342-y>
- The Digital Trade and Data Governance Hub** (2022) *The Participatory Indicator of Data Governance*. Global Data Governance Mapping. Available at <https://globaldatagovernancemapping.org/attributes/participatory> (accessed 15 December 2022).
- The GovLab** (2016) OPEN DATA’S IMPACT: Open data is changing the world in four ways. Open Data’s Impact – The GovLab. Available at <http://odimpact.org/> (accessed 15 December 2022).
- The GovLab** (2018) Data stewards: Data leadership to address the challenges of the 21st century. The GovLab. Available at <https://blog.thegovlab.org/archive/data-stewards-data-leadership-to-address-the-challenges-of-the-21st-century> (accessed 15 December 2022).
- The GovLab** (2021) The third wave of open data toolkit: Operational guidance on capturing the institutional and societal value of data re-use. The GovLab. Available at <https://files.thegovlab.org/The-Third-Wave-of-Open-Data-Toolkit.pdf> (accessed 15 December 2022).
- UN World Food Programme (WFP)** (2016) WFP Guide to Personal Data Protection and Privacy. Available at <https://docs.wfp.org/api/documents/e8d24e70cc11448383495caca154cb97/download/> (accessed 15 December 2022).
- UNCTAD** (2021) *Digital Economy Report 2021: Cross-Border Data Flows and Development - For Whom the Data Flow*. New York: United Nations Publications.
- UNICEF** (2020) *UNICEF Policy on Personal Data Protection*. UNICEF. Available at <https://www.unicef.org/supply/media/5356/file/Policy-on-personal-data-protection-July2020.pdf> (accessed 15 December 2022).
- UNICEF Division of Data, Research and Policy** (2015) Unicef procedure for ethical standards in research, evaluation, data collection and analysis. Available at <https://www.unicef.org/media/54796/file> (accessed 15 December 2022).
- United Nations** (2020) Data Strategy of the Secretary General for Action by Everyone, Everywhere. Available at [https://www.un.org/en/content/datastrategy/images/pdf/UN\\_SG\\_Data-Strategy.pdf](https://www.un.org/en/content/datastrategy/images/pdf/UN_SG_Data-Strategy.pdf) (accessed 15 December 2022).
- United Nations High Commissioner for Human Rights** (2018) A human rights-based approach to data. OHCHR. Available at <https://www.ohchr.org/sites/default/files/Documents/Issues/HRIndicators/GuidanceNoteonApproachtoData.pdf> (accessed 15 December 2022).
- United Nations High Commissioner for Refugees (UNHCR)** (2018) Guidance on the protection of personal data of persons of concern to UNHCR. Refworld. Available at <https://www.refworld.org/docid/5b360f4d4.html> (accessed 15 December 2022).
- Verhulst S** (2020) Unlock the hidden value of your data. *Harvard Business Review*. Available at <https://hbr.org/2020/05/unlock-the-hidden-value-of-your-data> (accessed 15 December 2022).
- Verhulst S** (2021a) *Data stewardship re-imagined | Capacities and competencies | Data stewards network*. Medium. Available at <https://medium.com/data-stewards-network/data-stewardship-re-imagined-capacities-and-competencies-d37a0ebaf0ee>.
- Verhulst S** (2021b) *Re-use of smart city data: The need to acquire a social license through data assemblies*. Medium. <https://medium.com/data-stewards-network/re-use-of-smart-city-data-the-need-to-acquire-a-social-license-through-data-assemblies-c096c1694cfc>.
- Verhulst, S. G.** (2023). Operationalizing digital self-determination. *Data & Policy*, 5. <https://doi.org/10.1017/dap.2023.11>
- World Bank** (2021) *World Development Report 2021: Data for Better Lives*. Washington, DC: World Bank.

**World Economic Forum (WEF)** (2021) Resetting Data Governance: Authorized Public Purpose Access and Society Criteria for Implementation of APPA Principles. Available at [https://www3.weforum.org/docs/WEF\\_Resetting\\_Data\\_Governance\\_2021.pdf](https://www3.weforum.org/docs/WEF_Resetting_Data_Governance_2021.pdf) (accessed 15 December 2022).

**World Health Organization (WHO)** (2020) Principles. WHO Data Principles. Available at <https://www.who.int/data/principles> (accessed 15 December 2022).