Data Governance Principles
Research & Development Project
A Report by the Centre for Intellectual Property and Information Technology Law (CIPIT), Strathmore University, Nairobi, Kenya.

Report Authors:
Amrit Singh P. Labhuram, Rachel Achieng’ Odhiambo,
Peacel Atim, Amanda Manyame
# Table of Contents

Executive Summary ........................................ ii
Introduction ................................................. 1
Objective ................................................... 1
Methodology ................................................ 1
   Validation Meetings ..................................... 2
Key Findings ............................................... 2
   Definitions and Concepts .............................. 2
   Data Governance Principles ......................... 4
   Data Governance Frameworks ....................... 7
   Data Governance in Africa ......................... 9
      Open data Initiatives in Africa .................. 9
      The Case of Kenya ................................ 10
   Data Governance within the Private Sector ...... 11
Challenges ............................................... 11
Concluding Remarks ..................................... 12
Digitalisation has resulted in increased data collection which forms an integral role in various industries and day-to-day activities thus the need to ensure that the data is managed in a manner that not only protects the integrity and security of that data but the rights and interests of the organisations processing the data and the people whose data may form part of the data sets being processed. Consequently, processes and standards have been developed to ensure the management or governance of that data. As a result of digitisation and the large volumes of data being processed in the digital era and the interdisciplinary nature in which processes and standards need to be applied, there has been an increase in these data governance processes. The interdisciplinary nature of data governance and the three interrelated concepts—Data (Information) Management, Enterprise Information Management and Data (Information) Architecture—resulted in different terms being used to refer to data governance. In this Report we use the term ‘data governance’.

The objective of this research is to provide a baseline understanding of what data governance is, the principles and frameworks applied in the governance of data and examine how these principles and frameworks translate in the African context. We found that data governance encompasses legal and human rights requirements, technological, security and economical considerations. These aspects inform, not only how data is managed within an organisation but determines the principles that are considered is that governance and the frameworks that are developed. In light of this data governance can essentially be defined as the use of authority combined with policy to ensure the proper management of data assets.

We record and provide a number of principles for data governance that can be used by research to focus on important data governance issues and practitioners to develop an effective data governance strategy and approach.1 We therefore incorporate data governance sources from practitioners.

---

Introduction

With the proliferation of digital technologies and the internet, data has become an asset. Data is not a mere commodity but an asset for driving businesses, creating value, and inspiring new ventures. Similar to financial gains, data should be treated sustainably and governed well, so as to give organisations competitive advantage. It is necessary to define how data is stored, archived, backed up, and protected. Lawyers are brought in to understand rules and regulations, which can and often do, frequently change. Moreover, a set of procedures need to be developed to define how data will be accessed by authorized personnel. Laws, standards and organisational policies have also been enacted and adopted to ensure that data, as an asset retains its integrity, is secure and human rights are protected. This data-driven governance is a new approach in which governance of data within an organisation is established at the data level and not at the system level.

Accordingly, data governance is the management of data within the possession of an organisation, that is, the management of cross-border data flows by national governments and corporations, as well as management of data that forms part of corporate data. The key focus areas of data governance include availability, usability, consistency, data integrity and data security. It also includes establishing processes to ensure effective data management. Data governance encompasses the people, processes, and information technology required to create a consistent and proper handling of data across an organisation or country. It provides all data management practices with the necessary foundation, strategy, and structure needed to ensure that data is managed as an asset and transformed into meaningful information.

It is the above considerations that inform the principles and frameworks that are used in practice by organisations to ensure compliance with regulatory requirements, international standards and overall ensuring data is managed effectively.

Objective

To contribute to the body of evidence available for those influencing policy in data protection, data bias, open data, and other issues pertaining to data governance, with a focus on issues relevant to the Global South, in particular Africa.

Methodology

This report examines what data governance is and provides introductory and background information taking into account the global and African contexts. This report does not purport to be a definitive representation of all the relevant standards and policies that are currently being applied. Notwithstanding, a literature review was done to collate studies and information on data governance that has already been conducted. The research was conducted to examine the following issues:

- What data governance is and why it matters?
- Who is impacted by data governance especially from an African perspective?
- What data governance principles are being applied presently? How do these
Validation Meetings
To provide understanding of data governance from both theory and practice, we identified a representative mix of stakeholders from geographically diverse regions on the continent. The stakeholders included personnel and organisations that could affect or be affected by the data governance process. These included:
- Institutional researchers, data managers, data architects, and business intelligence staff
- IT or system managers
- Relevant personnel from Tech companies including: The Facebook Group of Companies (WhatsApp, Instagram, Facebook), TiKTok, Twitter, and Google
- Data Protection Authorities and relevant ICT regulatory authorities and bodies

We focused on completing the stakeholder consultation activities through workshop sessions and reviews after each session. In the sessions we began by presenting the research conducted so as to establish an acceptable baseline of knowledge within the stakeholder group. We then conducted the sessions in a way that allowed participants to use that knowledge to provide practical feedback, based on their respective experiences. We then prioritised those ideas and feedback between the group using discussion and heat-map voting. We used a multi-stakeholder sectoral approach and experts across many sectors in order to test, and confirm or negate our research assumptions and ultimately provide well-evidenced data governance principles and data governance framework guidelines that when implemented will contribute positively.

Key Findings
Definitions and Concepts
Data is a set of characters, which have no meaning unless seen in the context of usage. Most scientific sources use the terms “information” and “data” interchangeably. In essence data governance ensures that data and information are managed appropriately. Theoretically, data governance describes the process and defines responsibilities. Data managers then work within its framework.

To understand what data governance is, it is important to understand that there are essentially three interrelated concepts—Data (Information) Management, Enterprise Information Management and Data (Information) Architecture—in terms of which managing data assets in a formal manner has been established.\(^5\)


\(^6\)Enterprise Information Management is the program that manages enterprise information assets to support the business and improve value. Enterprise Information Management manages the plans, policies, principles, frameworks, technologies, organizations, people, and processes in an enterprise toward the goal of maximizing the investment in data and content - DAMA International, DAMA-DMBOK: Data Management Body of Knowledge, 2nd ed. (Basking Ridge, NJ: Technics Publications, 2017).

\(^7\)A master set of data models and design approaches identifying the strategic data requirements and the components of data management solutions, usually at an enterprise level. - DAMA International, DAMA-DMBOK: Data Management Body of Knowledge, 2nd ed. (Basking Ridge, NJ: Technics Publications, 2017).

\(^8\)<https://learning.oreilly.com/library/view/data-governance/9780124158290/xhtml/CHP002.html#S0010tit>
Data governance is about the roles, responsibilities, and processes for ensuring accountability for and ownership of data assets. Data flows occur at global, national, and organisational level. These data assets include the data held in national ID systems by governments, KYC data held by organisations and personal data being processed by social media platforms. Accordingly, there are several reasons to manage data flows including:

- to safeguard the privacy of data subjects
- to meet other regulatory objectives, such as access to information for auditing purposes;
- for national/ organisational and cybersecurity reasons; and
- development of domestic capacity in data-intensive sectors, as a form of digital industrial policy.

As aforementioned, data governance has to do with the management of data flows at different levels. Accordingly, there are different aspects of data governance affecting or having an impact on different stakeholders. These are:

<table>
<thead>
<tr>
<th>Aspect of Data Governance</th>
<th>Focus</th>
<th>Affected Stakeholders</th>
</tr>
</thead>
<tbody>
<tr>
<td>Technology</td>
<td>Developing standards, apps and services for data management</td>
<td>Standardisation bodies (ITU, IETF, ISO, IEEE), Internet industry, software developers, academia</td>
</tr>
<tr>
<td>Economy</td>
<td>Use of data as the basis for the Internet business model</td>
<td>Internet companies, business associations, trade policy community</td>
</tr>
<tr>
<td>Security</td>
<td>Use of data by governments for protection of national security and the fight against crime</td>
<td>Security services, law enforcement agencies, Interpol, UNODC</td>
</tr>
<tr>
<td>Law and Human Rights</td>
<td>Protection of privacy, jurisdiction in transboundary data cases</td>
<td>Civil society, UN Human Rights Council, courts, academia</td>
</tr>
</tbody>
</table>

The information in the Table was extracted from DigWatch[^1^]

[^1^](https://dig.watch/issues/data-governance)
As a result of the different concepts at play, there is no single definition; however, it can be said that data governance involves the use of authority combined with policy to ensure the proper management of data assets.\(^\text{10}\)

### Data Governance Principles

Principles are a set of statements that describe the basic doctrines of data governance. A robust principle should enable good quality decisions to be made and enforceable policies and standards to be created.\(^\text{11}\)

Researchers have proposed initial frameworks for data governance\(^\text{12}\) and have analyzed influencing factors\(^\text{13}\) as well as the morphology of data governance.\(^\text{14}\) A number of data governance principles have emerged. These principles can be grouped into four categories — organisation, alignment, compliance and common understanding. Data governance principles should thus reflect and preserve the value to society from the sharing and analysis of anonymised datasets as a collective resource.\(^\text{15}\)

---


Data Governance Principles and Research and Development Project

Data Governance

- Organisation
  - Decision Rights
  - Balanced roles
  - Separation of Duties and Concern
  - Stewardship
  - Ownership
  - Improved Coordination of Decision Making

- Alignment
  - Meeting Business Needs
  - Aligning Business and IT
  - Developing Data Strategy
  - Defining Data Quality Requirements
  - Reducing Errors of Use
  - Effective Policies and Procedures

- Compliance
  - Accountability
  - Policy Enforcement
  - Openness
  - Privacy
  - Due Diligence
  - Security
  - Data Quality

- Common Understanding
  - Shared Data Commons
  - Use of Standards
  - Metadata Management
  - Standardised Data Models
  - Standardised Operations
  - Facilitates Communication

Image illustrating the breakdown of the four categories extracted from Brous P., et al\textsuperscript{16}

The four categories of data governance principles are as follows:

**Organisation**

Organisational goals are the first consideration when developing data governance frameworks. It is also important to take into consideration the organisational dimensions and specify the framework for decision rights and accountabilities to encourage desirable behaviours for the use of data.

These organisational goals also help to measure an organisation’s performance. Organisations need to consider their functional goals and organisational form. That is, consider the tasks needed to be fulfilled. Organisational form, looks at the structure in which responsibilities are specified and assigned and the organisational processes. In terms of this, the data governance model consists of roles, decision areas, main activities and responsibilities. There is a need to establish clear communications and patterns that would aid in handling policies for quick resolution of issues.

**Alignment**

Data governance should ensure that data meets the needs of the business. A data governance program must be able to demonstrate business value. Business uses for data establish the extent to which specific policies are appropriate for data management. Data governance should ensure data is useful and that the information should be helpful to its intended users, or support the usefulness of other disseminated information. While governments want to achieve the goals of data governance, they often have difficulty justifying the effort unless it has a practical concrete impact on business. Data governance also provides the framework for addressing issues such as improving data quality. Data quality is necessary to ensure that data management activities are in line with the overall business strategy. The strategy should include the strategic objectives which are pursued by data quality management and how it is aligned with the company’s strategic business goals and overall function scope.

**Compliance Monitoring and Enforcement**

Data governance includes clearly defining the authority to create and enforce data policies. It is thus imperative that an organisation establishes and enforces policies and processes around the management of data within the organisation. Organisational personnel such as the information technology and business teams therefore need to develop and determine a framework

---

of policies which are applicable across the whole organisation.\textsuperscript{26} The considerations for the framework of policies include, safeguards for the data sets, legal responsibility and accountability,\textsuperscript{26} as well as mechanisms to ensure that the organisation meets its legally defined obligations.\textsuperscript{27}

**Common Understanding**

Governing data appropriately is only possible if it is properly if the purpose and uses of the data is understood. It is therefore important to develop an enterprise data model,\textsuperscript{28} that way an organisation is able to understand its data uses and purposes over time, even when these uses and purposes change.\textsuperscript{29} Thus a common understanding of the purposes and uses of data need to be established across the organization.

In conclusion, data governance comprises legal and human rights requirements, technological, security and economical considerations. It is these aspects that inform how data is managed within an organisation, and determine the principles that are considered in the development of frameworks. The data governance principles can be grouped into four categories — organisation, alignment, compliance and common understanding. Notwithstanding, data governance is not a one size fits all, according to these four categories are used to inform the principles that an organisation will apply depending on its individualistic needs.

**Data Governance Frameworks**

Data governance may best be thought of as a function that supports an organization’s overarching data management strategy. Such a framework provides an organization with a holistic approach to collecting, managing, securing, and storing data.


\textsuperscript{26}Tallon, P.P.: Corporate governance of big data: perspectives on value, risk, and cost. Computer 46(6), 32–38 (2013)


The table below records some of the broad considerations that are currently being applied and informs data governance frameworks that are eventually adopted by an organization.

<table>
<thead>
<tr>
<th>Data Governance Considerations</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Value &amp; outcomes</td>
<td>That is, aligning data governance with a business or national goal or outcome.</td>
</tr>
<tr>
<td>Data governance strategies</td>
<td>Consider the objectives, principles and groups for a new or newly aligned data governance program.</td>
</tr>
<tr>
<td>Methods</td>
<td>Consider the people, processes and technologies that will be affected by data governance strategies applied by the organization.</td>
</tr>
<tr>
<td>Data management structures &amp; technologies</td>
<td>Consider the underlying concepts and technologies that can help establish and enforce data governance at the application or data level, within the organization.</td>
</tr>
</tbody>
</table>
| Accountability & decision rights               | Make sure your team is on board with your data governance strategy, and hold the chief data officer, data stewards, data owners, and employees accountable.  
  Clearly define who can make which decisions about your data. |
| Trust                                          | Consider the following questions:  
  ● Can you trust all of your data sources?  
  ● Are you in control of your data throughout its lifetime?  
  A trust model of data governance builds in mechanics to account for a distributed data ecosystem, and you should strive to understand the history and lineage of your data inputs so you can manage expectations and results. |
| Transparency & ethics                          | Keep data analytics about data governance — open and visible, with clear established decision-making processes, so it would stand up to external scrutiny. |
| Risk & Security                                | Risk and security are two major drivers for organizations to engage in data governance. These should be considered in addition to any business outcomes. |
| Education & Training                          | Consider the following questions:  
  ● Do you have a program to train employees on the basic principles of data governance?  
  ● Do you train new data stewards?  
  Establish an ongoing training program to keep data governance in focus. |
Data governance frameworks are determined by the uses of data, processing and storage policies adopted in the organisation, legal and best practise obligations, as well as other considerations that may be unique to an organisation. Accordingly, data governance is not a one size fits all and so the frameworks adopted by organisations are specific to the organisation and its need. Notwithstanding, the key considerations — value & outcomes, data governance strategies, methods, data management structures and technologies, accountability and decision rights, trust, transparency and ethics, risk & security and education and training — have a bearing on the frameworks developed and adopted by the organisation. It is by considering these that an organisation determines the parameters of their framework.

Data Governance in Africa

Open data Initiatives in Africa

Governments decide, through their own political processes, how the digital ecosystem under their jurisdiction will be governed to produce the social, economic and political outcomes that suit their contexts. The main objective of these initiatives is to increase transparency in government dealings.30

The Africa Open Data Network was created to strengthen efforts to increase the provision and use of open data in the region.31 The objectives of the Network were to increase awareness of the value of open data in Africa; support and scale activities of open data users in Africa; and promote the adaptation and contextualization of open data principles and practices.32

Many open data initiatives in Africa have struggled to find solid ground or have been deemed unsustainable by failing to show evidence of value and impact to the citizens who should be the primary beneficiaries.33 Open data initiatives thrive when there are “proper infrastructures, a high technology literacy rate, adapted national policies and strategies, national leadership, local intermediaries, local competencies and plenty enthusiasm among public institutions, civil societies, ICT companies, non-governmental organisations (NGOs) and academics” which are lacking in Africa.34

Most African countries lack any kind of comprehensive guidelines, technical standards, and management procedures for their (open) data.35 “Open data is not yet entrenched in law in the continent, and the legal frameworks supporting it are either incomplete or directly absent. Implementation and resourcing are also very weak.”36 Most African countries require technical and financial assistance to start or support open data initiatives. Some of the initiatives are usually vague and lack high-level

33https://www.stateofopendata.od4d.net/chapters/regions/sub-saharan-africa.html
34https://www.stateofopendata.od4d.net/chapters/regions/sub-saharan-africa.html
political backing.\textsuperscript{37} Morocco was the first country in Africa to launch an open data initiative.

The Constitution of Kenya 2010 grants its citizens the right of access to information.\textsuperscript{38} This right empowers the people to have access to information held by the State and mandates the State to publish and publicize any information that affects the country.\textsuperscript{39} In pursuance of this, a Kenyan civil society group, in 2011, called on the government to release financial data in order for citizens to examine how the government managed public resources.\textsuperscript{40} To that extent, the then President launched the Kenyan Open Data Initiative (KODI) to enable citizens to access information held by the government. The first set of data published on the portal was the 2009 census, national and regional expenditure, and information on key public services.\textsuperscript{41}

However, this initiative had some challenges. The first challenge was that the government bodies rarely shared the information they had with the citizens let alone other government bodies.\textsuperscript{42} The other challenge was that some institutions were corrupt and kept information secret in order to minimise the data shared so that they could advance their own interests. This corruption led to resistance against efforts to release information for accountability purposes.\textsuperscript{43} Despite these challenges, KODI was beneficial when it was in use. When the initiative was first launched in July 2011, “it held more than 160 datasets organised under six subheadings: education, energy, health, population, poverty and water and sanitation. The platform included newly created geospatial boundaries for Kenya’s 47 counties and geocoded datasets ... Data was pulled in from the national census and government ministries as well as from the World Bank.”\textsuperscript{44} “There were over 200 datasets released at the time of launch. As of June 2016, there were 849 datasets that had been uploaded to the site. The portal received approximately 1.1 million unique visits in 2013, while more than 5,500 datasets were downloaded and embedded into various websites and blogs.”\textsuperscript{45}


\textsuperscript{38} Article 35, Constitution of Kenya (2010).

\textsuperscript{39} Article 35, Constitution of Kenya (2010).

\textsuperscript{40} Centre for Public Impact, The Kenyan Open Data Initiative, 2016 \url{https://www.centreforpublicimpact.org/case-study/open-data-kenya}.

\textsuperscript{41} Centre for Public Impact, The Kenyan Open Data Initiative, 2016 \url{https://www.centreforpublicimpact.org/case-study/open-data-kenya}.

\textsuperscript{42} Centre for Public Impact, The Kenyan Open Data Initiative, 2016 \url{https://www.centreforpublicimpact.org/case-study/open-data-kenya}.

\textsuperscript{43} Centre for Public Impact, The Kenyan Open Data Initiative, 2016 \url{https://www.centreforpublicimpact.org/case-study/open-data-kenya}.

\textsuperscript{44} Centre for Public Impact, The Kenyan Open Data Initiative, 2016 \url{https://www.centreforpublicimpact.org/case-study/open-data-kenya}.

Data Governance within the Private Sector

Digitalisation has and continues to transform economies and societies globally. In Africa, the digital economy has promoted access to a number of services such as financial services and access to information, that were not so easily accessible. It has, however, highlighted challenges such as the digital divide and affordability to access the internet.\(^\text{46}\) Notwithstanding, data has become increasingly important for organisations. Accordingly, data governance concerns all sectors of the economy encompassing social, cultural, technical, and legal issues. Data flows across borders are integral to the global digital economy and a necessary input for reaping the benefits of digitalisation. Appropriate governance and safeguards for the collection, processing and storage of this data are becoming increasingly important.

Data governance within private organisations is worth noting because of the large volumes of data that these organisations collect, process and store. Within an organisation there are several aspects of data management, for example, designing, storage of data, ensuring the quality of the data, ensuring compliance with regulation and protection of user’s privacy.\(^\text{47}\) Multinational organisations and the policy influencers are faced with considerations of cross-border transfer of data, which is particularly challenging in Africa because of a patchwork of regulations across the continent. This makes it difficult to implement data governance frameworks across the organisations. This has resulted in multinational organisations applying EU regulations and standards, or using these as a set standard. In addition to the challenges of regulation, political will by countries to implement regulations that will afford, particularly multinational organisations, some form of standardisation when it comes to the collection, storage, processing and cross-border transfer of organisational data, is another challenge.

Challenges

In establishing a data governance framework, several challenges have been encountered by organisations. Since data governance frameworks are informed by the practices of the organisation at larger, the involvement of different departments is necessary. Additionally, there is a need for the organisation to be flexible and willing to pivot where necessary. A data governance framework should not impede work efforts, instead it should make work easier.\(^\text{48}\) Further, challenges with determining the technological tool to be used to manage can be difficult and to use.

Some of the challenges identified include:

- Concerns that there is too strong encryption of data by industry and users which can limit access to data for security reasons.
- Concerns that there is a loss of trust by users if data is shared with governments without legal constraints, increased privacy protection may reduce use of data and profit of Internet companies.
- Weak administrative data systems.
- Legal systems, policies and frameworks are not keeping up with the fast-changing data ecosystem. Adequate regulations to govern the rapidly evolving modern data ecosystem are lacking. Legal frameworks


\(^{48}\) [https://www.varonis.com/blog/data-governance/](https://www.varonis.com/blog/data-governance/)
regulating data sharing, security and coordination are still limited in most African countries.

- Lack of incentives and motivation to follow new data governance initiatives.
- Digital divide and poor data literacy (lack of simplicity for users). Most people in Africa lack access to ICTs. In the latest Telecommunications Infrastructure Index (TII), Africa was ranked worst in terms of regional leverage and access to ICT. The lack of access to ICTs and low levels of data literacy is also evident with public government officials. This limits the use of data, data-policy design and implementation and consequently, data governance.
- Difficulty in selecting technology and tools to manage a data governance strategy.
- Lack of flexibility in data governance efforts to suite particular teams needs
- “Most African countries lack any kind of comprehensive guidelines, technical standards, and management procedures for their (open) data. Open data is not yet entrenched in law in the continent, and the legal frameworks supporting it are either incomplete or directly absent. Implementation and resourcing are also very weak.”

More needs to be done when it comes to data governance principles and the considerations that inform the frameworks, especially when being developed and applied in Africa. This is so because of the, aforementioned, challenges around the regulations of cross-border data as well as patchwork in the regulation of the digital space on the continent. In addition there is a need to separately regulate the processing of personal data and other data, which would align with the way in which data as a whole is amalgamated. Further, on the continent, there is a need to revisit and identify the gaps and challenges arising from the gaps of the laws and regulations that predate most of the technological advancements that have resulted in the use of large volumes of data.

Concluding Remarks

Data is not a mere commodity but an asset for driving businesses, creating value, and inspiring new ventures. It is necessary to define how data is stored, archived, backed up, and protected. Procedures need to be developed to define how data will be collected, stored and processed by authorized personnel. Laws, standards and organisational policies have also been enacted and adopted to ensure that data, as an asset retains its integrity, is secure and human rights are protected. This data-driven governance is a new approach in which governance of data within an organisation is established at the data level and not at the system level. Accordingly, data governance is the management of data within the possession of an organisation\textsuperscript{50}, that is, the management of cross-border data flows by national governments and corporations, as well as management of data that forms part of corporate data.\textsuperscript{51} Data governance comprises the people, processes, and information technology required to create a consistent and proper handling of data across an organisation or country. It is from these considerations that data governance principles and frameworks are developed and implemented


\textsuperscript{50} The Data Governance Institute, <http://datagovernance.com/defining-data-governance/>

in practice by organisations to ensure compliance with regulatory requirements, international standards and overall ensuring data is managed effectively.

On the continent data governance is not without its challenges and sometimes unique considerations. There are challenges around the regulations of cross-border data as well as patchwork in the regulation of the digital space on the continent. In addition there is a need to separately regulate the processing of personal data and other data, which would align with the way in which data as a whole is amalgamated. Further, on the continent, there is a need to revisit and identify the gaps and challenges arising from the gaps of the laws and regulations that predate most of the technological advancements that have resulted in the use of large volumes of data.

Notwithstanding, data governance is not a one size fits all and it is important to bear the needs of the organisation, human rights, legal provisions and the best tools to implement decided upon data governance frameworks in mind.
This study was made possible by a grant provided by the Hewlett Foundation. We thank the organization for their continued support.