



# NATIONAL METADATA MANAGEMENT GUIDELINE



# Table of Contents

<b>Acknowledgment .....</b>	<b>2</b>
<b>1. Introduction .....</b>	<b>3</b>
1.1 Definition, Purpose and Scope .....	3
1.1.1 Definition .....	3
1.1.2 Purpose .....	3
1.1.3 Scope and applicability .....	4
<b>2. Understanding Metadata .....</b>	<b>5</b>
2.1 Benefits of Metadata Management .....	5
2.2 Types of Metadata .....	5
2.3 Metadata Standards .....	6
2.4 Metadata lifecycle management .....	6
<b>3. Metadata Principles, Development Guidelines, and Roles &amp; Responsibilities .....</b>	<b>8</b>
3.1 Guiding Principles .....	8
3.2 Metadata development guidelines .....	9
3.3 Mapping of Roles and Responsibilities for National Metadata Management .....	9
<b>4. Metadata Requirements .....</b>	<b>10</b>
4.1 Minimum Metadata elements applicable to all datasets .....	10
4.2 Additional Metadata for Administrative Data .....	11
4.3 Additional Metadata for Statistical Data .....	11
4.4 Additional Metadata for Geospatial and Big Data .....	12
<b>Annex 1: National Metadata Template for Public Institutions .....</b>	<b>13</b>
Section A: Core National Metadata template (All Datasets) .....	13
Section B: Administrative Data (if applicable) .....	14
Section C: Survey/Census Data (if applicable) .....	14
Section D: Geospatial and Big Data (if applicable) .....	15
<b>Annex 2. Glossary of Metadata Terminology .....</b>	<b>16</b>

# Acknowledgment

The development of the National Metadata Management Guidelines builds on the National Data Sharing Policy and the National Data Governance Framework. These guidelines establish a unified approach for documenting, organizing, and managing metadata across public institutions. By promoting consistent standards and practices, they strengthen the clarity, traceability, and usability of data needed to support national planning, digital transformation, and evidence-based decision-making.

The National Institute of Statistics of Rwanda (NISR) led the development of these guidelines in close collaboration with the Rwanda Information Society Authority (RISA), the Ministry of ICT and Innovation (MINICT), and technical partners, including Cenfri. Their collaboration ensured that the guidelines align with national priorities for interoperability, the National Data Infrastructure, and the long-term vision of creating a seamless and trusted data ecosystem.

NISR extends its appreciation to members of the National Statistical System (NSS), sector institutions, and all stakeholders whose contributions and feedback enriched the development of these guidelines and ensured their practicality and relevance.

The publication of the National Metadata Management Guidelines represents an important step toward strengthening data governance, enhancing data quality, and ensuring that metadata serves as a foundation for effective data sharing, integration, and innovation. These guidelines reaffirm Rwanda's commitment to responsibly managing data as a national asset for development.

# 1. Introduction

Metadata serves as the backbone of effective data management and utilization, providing the essential context, meaning, and structure that make data understandable, comparable, and reusable across institutions and sectors.

In fulfillment of this principle, the National Institute of Statistics of Rwanda—mandated by the Government of Rwanda through the National Data Sharing Policy approved on May 26, 2025—issues these guidelines to ensure standardized production, documentation, sharing, and storage of metadata across the National Statistical System (NSS). These guidelines promote the generation of high-quality, discoverable, interoperable, and reusable data by all official data-producing entities in Rwanda, in accordance with Statistics Law No. 45/2013.

These guidelines establish a coherent framework for capturing, maintaining, and disseminating metadata throughout the data lifecycle. They address persistent challenges including fragmented documentation practices, limited interoperability between institutions, and inconsistent standards that undermine data quality and usability. By fostering consistency and integration, these guidelines will strengthen data comparability across institutions and empower policymakers, researchers, businesses, and citizens to make well-informed decisions with confidence.

## 1.1. Definition, Purpose and Scope

### 1.1.1. Definition

Metadata refers to the descriptive, structural, technical, and contextual information that gives meaning to data. It enables users to identify, locate, and understand datasets by explaining how data are created, collected, managed, and processed.

### 1.1.2 Purpose

The purpose of these guidelines is to establish systematic, standardized practices for metadata development and management across all Government of Rwanda (GoR)<sup>1</sup> entities, ensuring that datasets are well-documented, discoverable, accessible, interoperable, and reusable.

Specifically, the guidelines aim to:

- Define standardized requirements for metadata elements and documentation practices.
- Provide processes, tools, and templates for creating, storing, updating, and disseminating metadata.
- Clarify roles and responsibilities to ensure accountability in metadata governance.
- Strengthen transparency and trust to ensure data users understand the origin, methodology, and limitations of datasets.

---

<sup>1</sup>In this framework, “GoR” refers to all public institutions and government-affiliated entities in Rwanda, including ministries, departments, agencies, local governments, and state-owned enterprises.

- Align metadata practices with legal and ethical frameworks, including Rwanda’s Data Protection and Privacy Law (N° 058/2021) and international best practice.
- Encourage collaboration among MDAs to reduce duplication and improve efficiency.
- 
- Promote the Reuse and integration of data appropriately in new contexts.

### *1.1.3 Scope and applicability*

These guidelines apply to all GoR entities and cover all forms of data—structured, unstructured, administrative, and survey-based.

## 2. Understanding Metadata

### 2.1 Benefits of Metadata Management

Effective metadata management provides multiple benefits for public institutions and the wider data ecosystem:

- **Improved Discoverability:** Well-defined metadata makes datasets easier to locate, access, and reuse across government systems.
- **Enhanced Interoperability:** Standardized, structured metadata enables seamless data integration and exchange between institutions, reducing duplication and fragmentation.
- **Increased Transparency and Trust:** Comprehensive documentation of methodologies, sources and scope strengthens credibility, accountability and public confidence in data.
- **Regulatory and Standards Compliance:** Metadata aligned with national laws, regulations, and international standards ensures responsible, ethical and legally compliant data management.
- **Informed Decision-Making:** High-quality metadata allows policymakers, researchers, and stakeholders to accurately interpret and apply data for evidence-based decisions.
- **Operational Efficiency and Cost Savings:** Streamlined metadata practices minimize redundancy of data, promote reusability, and optimize institutional workflows.
- **Improved Data Quality:** Metadata supports the assessment of completeness, accuracy, and reliability of datasets, enhancing overall data quality.

### 2.2 Types of Metadata

Metadata can be categorized into several core types, each serving a specific role in managing and using data:

- **Descriptive Metadata:** Identifies, explains, and facilitates the discovery of datasets (e.g., title, abstract, keywords). Helps users quickly understand the dataset's scope and locate it efficiently.
- **Structural Metadata:** Describes how data is organized and how components relate (e.g., schemas, table relationships). Supports integration, processing, and storage.
- **Administrative Metadata:** Provides information on ownership, access rights, and retention policies (e.g., data owner, license type, access restrictions). Ensures accountability, governance, and compliance.
- **Technical Metadata:** Captures technical characteristics of datasets and storage systems (e.g., file formats, encoding, storage locations, APIs). Enables IT teams to manage and integrate data efficiently.
- **Provenance Metadata:** Tracks the origin, transformations, and versioning of data throughout its lifecycle (e.g., source system, transformation rules, version control). Ensures traceability, transparency, and data integrity.
- **Business Metadata:** Explains the meaning, purpose, and rules for interpreting data in operational or policy contexts (e.g., glossary terms, classifications, usage rules). Helps stakeholders apply data correctly.
- **Statistical Metadata:** Provides methodological and quality-related information specific to statistical datasets (e.g., sampling methods, indicators, aggregation levels). Supports researchers and statisticians in assessing reliability and suitability.

## 2.3 Metadata Standards

To ensure consistency, interoperability, and usability of data across GoR entities, metadata should align with recognized international standards. Key standards include:

- **DDI (Data Documentation Initiative):** Standard for documenting surveys and microdata throughout their lifecycle — from collection to archiving — ensuring reproducibility and long-term usability.
- **SDMX (Statistical Data and Metadata Exchange):** Standard for sharing and comparing aggregated data between agencies and countries in a structured way.
- **ISO 19115:** Standard for geospatial metadata, enabling integration with GIS systems and international interoperability.
- **DCAT (Data Catalog Vocabulary):** Improves dataset discoverability on open data portals by standardizing descriptions.
- **Dublin Core:** Lightweight, simple standard for cataloguing datasets and enabling basic search.
- **GSIM (Generic Statistical Information Model):** Provides a common framework for describing statistical information objects and workflows.
- **GSBPM (Generic Statistical Business Process Model):** Describes and standardizes the steps in statistical production processes.

**Application note:** The national metadata template will be based on these standards to ensure that all datasets documented across the GoR entities are interoperable, comparable, and reusable. All public institutions are required to systematically apply the template and reference relevant standards where appropriate, thereby maintaining national coherence while allowing flexibility for specialized data needs.

## 2.4 Metadata lifecycle management

Effective metadata management requires capturing, maintaining, and updating metadata throughout the entire data lifecycle. Metadata should remain accurate, interoperable, secure, and accessible from the point of data creation to long-term preservation. The following lifecycle approach supports metadata management:

- **Creation:** Metadata should occur as soon as data is collected or generated using national standardized template aligned with recognized standards. Essential elements, including dataset title, creator, description, format, contributor, dates, unique identifiers, source, and rights information, should be captured comprehensively to facilitate later discovery and management.
- **Storage:** Metadata should be centralized and secured in a data catalog, with searchable repositories that support version control, indexing, and interoperability with other systems. (Where metadata platforms are not yet in place, institutions may initially use Excel or similar simple tools.)
- **Maintenance and Quality Control:** Metadata should be updated whenever datasets are modified, whether through new variables, methodological changes, or updates to ownership. Automated validation tools, periodic quality audits, and transparent change logs help maintain accuracy, consistency, and completeness. (Any update must be addressed to NISR for Approval)
- **Sharing and Access:** Clear policies should be developed to clarify who can access specific metadata elements, and through which mechanisms it can be accessed such as APIs, portals, or data catalogues, etc. Openness should be encouraged where possible, with restrictions applied only for legal, privacy, or security reasons.

- **Security and Compliance:** metadata should be classified according to the sensitivity of the underlying data and institutional policy, with role-based access controls and encryption applied as needed. Adherence to Rwanda's Data Protection & Privacy Law and other regulatory frameworks ensures legal compliance and safeguards data integrity.
- **Capacity Building:** Regular training on standards, tools, quality controls, and security obligations should be provided to ensure that staff are able to manage metadata efficiently and stay current with emerging technologies and best practices.
- **Monitoring and Review:** key performance indicators should be developed by the NISR. Periodic reviews should be conducted to allow institutions to incorporate feedback, audit findings, and technological advances into metadata processes, ensuring continuous improvement.
- **Archiving and Preservation:** Apply secure disposal procedures when datasets are no longer required, following national retention and disposal guidelines. Securely handle any sensitive metadata, and keep only a non-sensitive catalog record if needed.



## 3. Metadata Principles, Development Guidelines, and Roles & Responsibilities

### 3.1 Guiding Principles

Effective metadata management is guided by principles that ensure transparency, consistency, interoperability, and long-term accessibility. These principles provide a foundation for institutions to manage metadata in ways that maximize its value for governance, integration, and reuse.

#### **Principle 1: Existence**

No dataset should exist without metadata. Metadata must be created and documented at the point of data collection or generation and maintained throughout the data lifecycle. National standardized templates and recognized frameworks should be applied to ensure completeness and comparability.

#### **Principle 2: Quality and Consistency**

Metadata must be accurate, comprehensive, and updated whenever the underlying dataset changes. Institutions should conduct regular quality checks, adopt controlled vocabularies, and provide mechanisms for users to flag inconsistencies or gaps. The national data quality and standards outline more details to ensure metadata quality.

#### **Principle 3: Interoperability and Integration**

Metadata should be structured using open standards and machine-readable formats to enable cross-institutional sharing, data exchange, and integration. A centralized and interoperable metadata catalogue should be maintained to support accessibility across government systems.

#### **Principle 4: Security and Privacy**

Because metadata may reveal sensitive information about datasets, institutions must apply classification schemes as outlined in national data classification and access guidelines. Staff should be trained to comply with metadata security and privacy policies.

#### **Principle 5: Roles and Responsibility**

Defined responsibilities are critical to effective metadata stewardship. Institutions must designate creators, reviewers, and custodians of metadata, supported by workflows for approval, version control, and audit trails.

#### **Principle 6: Sustainability and Improvement**

Metadata management should evolve with emerging technologies and user needs. Institutions should track key indicators such as completeness, timeliness, and usage. Regular reviews, staff training, and adoption of modern tools will ensure metadata remains relevant and effective over time.

#### **Principle 7: Preservation**

Metadata must remain usable and accessible as long as the underlying data is preserved. Persistent identifiers, archival processes aligned with retention policies, and proper documentation of standards ensure continuity and institutional memory.

## 3.2 Metadata development guidelines

Each institution is required to develop and maintain internal metadata aligned with the national metadata template, recognized standards, and the national governance framework. Internal metadata must be clear, practical, and consistently applied throughout the entire data lifecycle. At a minimum, institutions should:

- Define mandatory metadata elements to be captured for every dataset, ensuring completeness, comparability, and interoperability across institutions.
- Specify approval processes, version control, and update mechanisms to maintain metadata accuracy and integrity over time.
- Establish clear rules for metadata publication, access, and sharing, taking into account data sensitivity, security, and user needs.
- Integrate metadata management into institutional data governance structures, ensuring accountability and alignment with broader data management practices.
- Define roles and responsibilities, ensuring that staff at all levels understand their duties in creating, maintaining, and using metadata.
- Use approved tools and platforms to capture, store, and exchange metadata in line with national standards.
- Participate in capacity-building and quality assurance exercises to strengthen metadata practices and support continuous improvement.

## 3.3 Mapping of Roles and Responsibilities for National Metadata Management

To ensure effective metadata governance across GoR, the following roles and responsibilities shall be designated and fulfilled by all GoR entities.

S/N	Role	Key Responsibilities / Description	Where the Role Sits
1	Data Owner	<ul style="list-style-type: none"><li>• Approve and validate metadata content for datasets under their authority.</li><li>• Ensure compliance with organizational policies, legal frameworks, and metadata guidelines.</li><li>• Decide on dataset classification and access levels.</li></ul>	Institutional/ Sector level
2	Data Steward	<ul style="list-style-type: none"><li>• Ensure metadata quality, completeness, and consistency, aligned with the national metadata template and organizational glossaries.</li><li>• Monitor adherence to standards during metadata creation and updates.</li><li>• Collaborate with custodians, IT teams, and metadata managers to resolve inconsistencies or gaps.</li></ul>	Institutional/ Sector level
3	Data Custodian	<ul style="list-style-type: none"><li>• Manage storage, versioning, backups and technical infrastructure for metadata.</li><li>• Automate metadata capture and ingestion from systems.</li><li>• Ensure interoperability between metadata repositories.</li><li>• Implement schemas, APIs, and access controls.</li><li>• Monitor performance, security, and compliance.</li></ul>	Central (RISA) and Institutional

## 4. Metadata Requirements

Metadata requirements establish the minimum information that must accompany every dataset to ensure consistency, discoverability, interoperability, integration and usability across institutions. These requirements apply to all types of data (administrative, statistical, geospatial, and big data) and are aligned with international best practices as well as the national data governance framework.

### 4.1 Minimum Metadata elements applicable to all datasets

Every dataset must include a core set of metadata elements, regardless of its type, source, or origin. The following elements are mandatory:

- **Title:** The official name of the dataset, clearly describing its subject matter.
- **Description:** A concise summary of the dataset's content, scope, and purpose.  
This should include:
  - **Producer / Responsible Institution:** The government institution, ministry, agency, or department responsible for producing, maintaining, and updating the dataset.
  - **Target Population / Universe:** The group or entities about which information is collected (e.g., households, enterprises, land, livestock, individuals, etc.)
  - **Units of measurement:** The standard metric or scale used to quantify a variable (e.g., number of individuals, percentages, etc.)
  - **Keywords / Tags:** Controlled vocabulary terms and/or free-text keywords to improve search, discovery, and interoperability.
  - **Geographic Coverage:** The spatial extent of the dataset (e.g., national, regional, district, sector level).
  - **Reference period:** The specific point in time or span of time to which the data values relate. (e.g., Jan-Dec 2023)
  - **Frequency:** How often is data collected or updated (e.g., daily, monthly, quarterly, annually)
  - **Methodology / Process Documentation:** A brief explanation of how the dataset was created, including data sources, collection and computation method and processing steps.
  - **Format (File Extension / Machine-Readable Format):** The technical file format in which the dataset is provided (e.g., STATA, SPSS, CSV, Excel, JSON, API, PDF, etc.)
  - **Access Rights:** The level of accessibility (e.g., public, internal, confidential, restricted) and licensing or usage conditions.
  - **Product Type:** The type of dataset being provided (e.g., aggregated data, anonymized microdata, geospatial data).
  - **Version Information:** The dataset's version number, date of release, and details of any changes from previous versions.

## 4.2 Additional Metadata for Administrative Data

Administrative data is collected primarily for operational purposes within institutions but can be repurposed for statistical or analytical use. Metadata for such data must highlight its institutional, and operational context to ensure proper interpretation.

Minimum addition required metadata includes:

- **Data Management Procedures/System Name:** Describes the name of the system, platform, or workflow from which the administrative data originates. (e.g., CRVS, HMIS, EMIS, etc.,)
- **Reporting Units / Administrative Level:** Indicates the lowest administrative or organizational level that reports or collects the data within the system. (e.g., Health facilities, schools, district offices, etc.,)
- **Timeliness / Lag:** Refers to the typical time delay between the end of the reference period and the availability of the data. (e.g., Data available 30 days after the end of the reference month.)
- **Quality Notes:** Provides brief statements explaining any known data reliability issues or limitations, such as missing records, reporting delays, or changes in definitions.

## 4.3 Additional Metadata for Statistical Data

Statistical data, including survey and census data, requires methodological documentation to support reproducibility, comparability, and credibility.

Required metadata elements include:

- **Type of Survey/Study:** The general category of the survey or census. (census, household survey, enterprise survey, etc.) and its objectives.
- **Sampling Methodology:** Sampling frame, technique, and sample size and allocation.
- **Data collection Instruments:** Lists questionnaires and tools used to collect data
- **Disaggregation/Variables:** Shows main disaggregation dimensions available (sex, age, location, etc.)
- **Methodological Notes:** Data collection procedures, imputation, weighting, estimation methods.
- **Quality notes:** Provides short statements on known data issues, or limitations—such as accuracy, completeness, timeliness, or comparability.
- **Revision and Validation Process:** Procedures for data editing, validation, and updates.

## 4.4 Additional Metadata for Geospatial and Big Data

Geospatial datasets require specialized metadata elements to describe their spatial reference and usability.

- **Coordinate Reference System (CRS):** The spatial reference framework used (e.g., WGS84).
- **Geographic Extent:** Boundaries of coverage (national, regional, local, or specific coordinates).
- **Scale/Resolution:** The level of spatial detail (e.g., 1:50,000 maps, 30m raster).
- **Geospatial Data Format:** The format used to store spatial data. (e.g., Shapefile (.shp), GeoJSON, GeoTIFF)
- **Topology / Geometry Type:** The type of spatial features. (e.g., Point, line, polygon)
- **Accuracy and Quality Notes:** Positional accuracy, resolution limits, and error margins.
- **Data Capture Method/Data sources:** Method used to produce geospatial data. (e.g., Satellite imagery, GPS survey, digitized maps, administrative boundaries).

Big Data sources (social media, mobile, sensors, transactional data, etc.) introduce unique metadata needs due to volume, velocity, and ethical concerns.

- **Data Source type:** Platform, system, or technology from which the data originates. (Satellite imagery, social media feeds, or sensor generating the data, etc.,).
- **Acquisition Method:** How the data was accessed or obtained (APIs, partnerships, or web scraping, sensor network).
- **Data Volume:** The approximate size or number of records in the dataset (2 TB per month; 500 million records)
- **Algorithm / Analytical Method Used:** Description of models, scripts, or algorithms used to transform raw data.

## Annex 1: Rwanda National Metadata Template (GoR)

This template provides a standard structure for recording metadata across all datasets. Each institution should adopt and adapt it to fit its operational needs while ensuring alignment with the National Metadata Guidelines.

### Section A: Core National Metadata template (All Datasets)

Metadata Element	Description / Guidance	Example	Explanations/Importance
Dataset Title	Official name of the dataset	Household Consumption Survey 2024	Identifies and references the dataset clearly.
Description	A concise summary of the dataset's content, scope, and purpose. (1-3 Sentences)	This dataset provides information on household expenditures and living conditions in Rwanda.	Helps users quickly understand what the dataset is about.
Keywords/Tags	List of keywords to aid search and discovery	Birth, registration, marriage	Improves discoverability in catalogues or repositories.
Responsible Institution	Institution/department responsible for producing the dataset	Rwanda Revenue Authority	Establishes institutional ownership and accountability for the dataset.
Contact Person	Name, position, and email of the responsible officer	John Rwema, Senior Statistician, <a href="mailto:John.rwema@statistics.gov.rw">John.rwema@statistics.gov.rw</a>	Allows users to request clarifications, or support regarding the dataset.
Target Population / Universe	The group or entities about which information is collected	Households, enterprises, land parcels, livestock, individuals	Defines what the data represents, ensures correct interpretation and comparability
Geographical Coverage	Area covered by the dataset	National, Province, District, etc.	Gives the spatial scope and coverage of the data.
Reference Period	Time covered	January – December 2024	Provides the temporal context of the data, essential for trend analysis and relevance checks.
Frequency	How often data are collected or updated	Every quarter, daily,	Shows how often the data is produced, helping users plan for updates and monitor timeliness.
Format	Machine readable format (Extension)	Stata, SPSS, CSV, XLSX, XML, JASON, etc.	Ensures usability across tools
Units of Measurement	The standard metric or scale used to quantify a variable or data value.	Number of individuals, Rwandan Francs (RWF), percentages (%)	Ensures correct interpretation, comparability, and integration of data values.
Variable Definitions and Classifications	Standardized descriptions of key variables, codes, and classifications used.	(e.g., ISIC, ICD11,...)	Enables consistent interpretation and comparison
Data source	Origin of data	Survey, Census, Administrative record, etc.	Shows how the data was generated and its method
Access Rights	Public / Internal / Confidential / Restricted	Public use file	Defines access conditions for legal and ethical use
Version Information	Version number and release date	v1.0, released June 2025	Enables tracking of dataset updates and changes over time, ensuring correct citation and use.

## Section B: Administrative Data (if applicable)

When it is the administrative data, the following metadata element is filled as additional to the core metadata template in section A of the annex

Element	Description / Guidance	Example	Explanations/Importance
<b>Data Management Procedures/ Administrative System Name</b>	The name of the system/platform/workflow from which the data originates.	CRVS, HMIS, etc.	Identifies the source system, ensuring traceability and understanding of data origin
<b>Reporting Units / Administrative Level</b>	The lowest administrative level at which data is reported or collected/The organizational level that reports data into the system	Health facilities, Schools, District Offices, etc.	Clarifies who reports the data and at what level, aiding consistency and accountability
<b>Timeliness / Lag</b>	Typical time delay between the reference period and availability of data	Data available 30 days after the reference month	Shows data freshness and helps assess relevance and usability
<b>Quality Notes</b>	Quality notes for administrative data are brief explanations of data reliability and limitations, such as missing records, reporting delays, or changes in definitions.	-Some hospital reports were incomplete, affecting total patient counts. -Tax records from 2023 were delayed, so annual totals may be underestimated.	Alerts users to potential data issues, supporting proper interpretation and use

## Section C: Survey/Census Data (if applicable)

When it is the survey/census data, the following metadata element is filled as additional to the core metadata template in section A of the annex

Element	Description / Guidance	Example	Explanations/Importance
<b>Type of Survey/Study</b>	The general category of the survey or census.	Labour Force Survey (LFS), Demographic and Health Survey (DHS), etc.,	Helps users understand the purpose and nature of the data
<b>Sampling Methodology</b>	Frame, technique, sample size and allocation	Stratified random sampling, n=12,000 households	Explain how sample units are selected, ensuring representativeness.
<b>Data Collection Instruments</b>	Questionnaires, interview schedules, or tools used.	Structured household questionnaire (English & Kinyarwanda)	Clarifies how data was gathered, supporting interpretation
<b>Desaggregation</b>	Extent to which the data is disaggregated.	By Sex, Education, Age, Residence, etc.	Shows data depth for detailed analysis
<b>Quality Notes</b>	Quality notes explain a dataset's reliability and limitations, such as accuracy, completeness, timeliness, and comparability.	Poverty measurement changed in 2023, affecting comparability.	Alerts users to data strengths and weaknesses for informed use

## Section D: Geospatial and Big Data (if applicable)

Metadata Element	Description / Guidance	Example	Explanation / Importance (condensed)
<b>Geospatial</b>			
<b>Coordinate Reference System (CRS)</b>	The spatial reference framework used	WGS84	Ensures spatial data aligns correctly across systems
<b>Geographic Extent</b>	Boundaries of coverage (national, regional, local, or coordinates)	National, Kigali City, or specific coordinates	Shows the spatial coverage and area represented
<b>Scale / Resolution</b>	The level of spatial detail	1:50,000 maps, 30 m raster	Indicates how detailed or generalized the data is
<b>Geospatial Data Format</b>	The format used to store spatial data	Shapefile (.shp), GeoJSON, GeoTIFF	Ensures compatibility with GIS tools and systems
<b>Topology / Geometry Type</b>	The type of spatial features	Point, Line, Polygon	Describes the shape of spatial data for proper use
<b>Accuracy and Quality Notes</b>	Notes on positional accuracy, resolution limits, and error margins	±5 m accuracy; some boundaries approximate	Helps users assess reliability and precision
<b>Data Capture Method / Data Sources</b>	Method or source used to produce the geospatial data	Satellite imagery, GPS survey, digitized maps	Provides context on how data was produced and its reliability
<b>Big data</b>			
<b>Data Source Type</b>	Platform, system, or technology from which the data originates.	Satellite imagery, social media feeds, sensor networks, mobile phone call records	Identifies the nature and structure of the data.
<b>Acquisition Method</b>	How the data was accessed or obtained.	APIs, formal partnerships, web scraping, sensor network feeds	Shows how and under what conditions the data was obtained.
<b>Data Volume</b>	The approximate size or number of records in the dataset.	2 TB per month; 500 million records	Indicates storage and processing requirements.
<b>Algorithm / Analytical Method Used</b>	Models, scripts, or algorithms used to transform raw data.	NLP sentiment analysis model; machine learning classification script	Ensures transparency and reproducibility of results.



## Annex 2. Glossary of Metadata Terminology

Term	Definition
<b>Metadata</b>	Structured information that describes, explains, locates, or otherwise makes it easier to retrieve, use, or manage data. Metadata provides context for data, including its origin, format, structure, and meaning.
<b>Administrative Metadata</b>	Metadata provides information to help manage a resource, such as creation date, creator, version, access rights, and preservation details.
<b>Descriptive Metadata</b>	Information that describes the content, subject, and context of a resource, enabling discovery and identification (e.g., title, abstract, keywords).
<b>Structural Metadata</b>	Information about how a dataset or resource is organized internally, such as schema, relationships between data elements, and table structures.
<b>Statistical Metadata</b>	Metadata describing statistical data and processes, including methodology, sampling, classifications, variables, and indicators.
<b>Metadata Catalogue</b>	A centralized platform or system that stores and organizes metadata to enable users to search, access, and manage datasets across government institutions.
<b>Metadata Lifecycle</b>	The sequence of stages through which metadata passes—from planning and collection to dissemination, preservation, and disposal—ensuring it remains accurate and up-to-date.
<b>Data Steward</b>	An individual responsible for ensuring the quality, completeness, documentation, and compliance of data and metadata within an organization.
<b>Data Custodian</b>	The person or entity responsible for the safe custody, and storage of data, often with a technical or operational role.
<b>Interoperability</b>	The ability of systems, datasets, and metadata to be used and understood across different platforms and institutions, enabling seamless data exchange and integration.
<b>Provenance</b>	A detailed record of the origin, ownership, and transformations applied to a dataset, ensuring transparency and traceability.
<b>API (Application Programming Interface)</b>	A set of rules and protocols allowing systems or applications to access, query, and interact with datasets programmatically.
<b>Data Quality</b>	The degree to which data meets requirements for accuracy, completeness, reliability, relevance, and timeliness for its intended purpose.
<b>License</b>	The legal terms under which data may be used, shared, or redistributed. Examples include Creative Commons licenses and Open Data Commons licenses.
<b>Open Data</b>	Data that is freely available for anyone to use, reuse, and share, subject only to requirements that preserve provenance and openness.

