[March 2025]

Strategic Pathway 4

Data

*This **strategic pathway** establishes a geospatial data framework and custodianship guidelines for best practice collection and management of integrated geospatial information that is appropriate to ensure cross sector and multidisciplinary collaboration.*

*The **objective** is to enable data custodians to meet their data management, sharing and reuse obligations to government and the user community through the execution of well-defined data supply chains for organizing, planning, acquiring, aggregating, integrating, curating, analyzing, publishing and archiving geospatial information.*

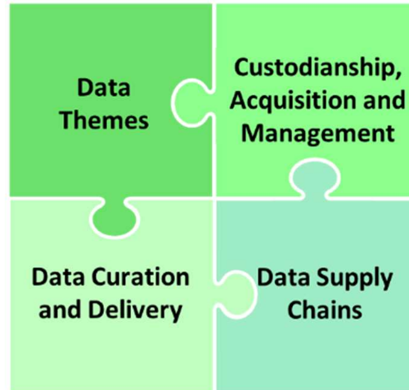
Summary

Geospatial data is the foundation on which governments base many decisions. It is used in policy development in the provision of government services. Its use is growing exponentially across all sectors for e-commerce, business intelligence, to make timely and accurate decisions, and to inform policy. Having access to the appropriate data and at the right time is crucial to good decision-making. It is data that provides new levels of insight into our past, present and future. For this reason, governments, businesses and the community need to know they are using the most appropriate and authoritative data for planning, analysis, navigation and visualization – good data underpins good decisions.

As the amount, variability and availability of data rapidly increases, the requirements for 'organized' geospatial data holdings have never been more important. Geospatial data has grown in use across almost every market and industry sector. Every part of government creates and consumes geospatial data. It is an important part of a country's 'digital currency', an asset that must be properly governed, designed and managed to provide enduring consistency, coherency and completeness in quality, accuracy, security and use. An ecosystem that fosters the proper collection, acquisition and management of geospatial data, leads to cutting-edge innovation and revolutionary methods across a range of sectors. Advances can already be seen in the health, financial, disaster management and transportation sectors; where geospatial data is enabling the discovery of new patterns and influences by combining geography, and social and cultural norms in a way that reveals new knowledge through enhanced visualizations, analysis and traceability.

This strategic pathway is primarily focused on supporting countries to build and maintain fundamental geospatial data themes. Nonetheless, many of the concepts and methodologies also apply to non-traditional forms of location-based data such as health data, as well as the development, use and uptake of "information" in administrative systems, such as land ownership, environmental land management. Geostatistical data integration is also explored in terms of the value created that can be created through interrelating geospatial and social, economic, and environmental data. The management and facilitation of data sharing between institutions is explained with reference to data supply chains. Data access systems and discoverability of data are explored in Strategic Pathway 5: Innovation along with the broader data ecosystem (e.g., via web services and enabling infrastructure for data sharing).

Common to all government and business applications are four key elements associated with data coordination that need to be achieved to enable an environment where innovation, and pioneering research and development can thrive. These four elements are:



- Data Themes – the organization of priority national data themes, aligned to the globally endorsed fundamental geospatial data themes.
- Custodianship, Acquisition and Management – leading to responsible collection, management, maintenance, use, dissemination and accountability of fit-for-purpose geospatial information.
- Data Supply Chains – and interlinkages that support cooperative data sharing and integration.
- Data Curation and Delivery – enables enduring accessibility and value of data, as an information resource for broader usage across all sectors.

These elements are underpinned by principles that promote consistent data management, sharing and reuse, so that data custodians may meet their obligations to government and the user community. These principles are put into practice through strategic actions that deliver and strengthen effective integrated geospatial information management. Tools, such as matrices, examples and checklists, are provided in the appendices to assist countries to work through concepts and processes to successfully complete each action. The overall structure for data is illustrated in and anchored by Figure 4.1.

When implemented, the actions (and their interrelated actions[1]) will enable the achievement of the four elements, which in turn will deliver significant and sustainable outcomes and benefits for a country. These outcomes include attaining:

- Increased range, scope and awareness of authoritative, integrated geospatial information available for decision-making and policy-setting to address economic, social and environmental challenges;
- Critical mass of coordinated data discovery to support national development and innovation, leading to economic growth and improved quality of life for citizens;
- Cost reduction through productivity improvements achieved via well-defined data supply chains that eliminate duplication and ensure standardized data is accessible to end users for integration and reuse;
- Ability to monitor and measure progress towards achieving broad socio-economic benefits, including the sustainable development goals, through access to quality geospatial information.

[1] Examples of the interrelated actions across Strategic Pathways are described in the introductory chapter; Solving the Puzzle: Understanding the Implementation Guide.

Elements of Data	Data Themes	Custodianship, Acquisition and Management	Data Supply Chains	Data Curation and Delivery

Guiding Principles

Data Governance	Metadata	Authoritative	Demand Driven
Consistent Identification	Standards	Timeliness	Efficiency
Data Quality Management	Accessibility	Provenance	Data security
	Reusable Formats	Integrity	Respected Rights

Key Actions for Strengthening Geospatial Information Management

Getting Organized
Data Framework
Data Inventory
Dataset Profiles

Capturing and Acquiring Data
Data Capture
Data Acquisition Program

Maintaining Accurate Positioning
Maintained Geodetic Infrastructure

Planning for the Future
Data Gap Analysis
Data Theme Road Map

Managing Data Sustainably
Data Custodianship Policy
Data Governance
Data Management
Maintained Metadata
Data Release
Data Storage and Retrieval Systems

Integrating Data
Geospatial and Statistical Integration
Geocoding and Aggregation
Data Supply Chains
Data Interoperability

Tools to Assist in Completing the Actions

- Global Fundamental Geospatial Data Themes
- Data Theme Description
- Data Inventory Questionnaire
- Dataset Profile Template
- Gap Analysis Matrix
- Data Theme Road Map template
- Principles for Data Custodianship Policy
- Data Governance Roles and Responsibilities
- Elements of a Data Management Plan
- Metadata Creation Checklist
- Data Release Guidelines
- Guidance for Improving Geodetic Infrastructure
- Global Statistical Geospatial Framework (GSGF)
- Guidance for Geospatial and Statistical Integration

Interrelated and/or Prerequisite Actions

- Geospatial Strategy (SP1)
- Country-level Action Plan (SP1)
- Governance Model (SP1)
- Geospatial Information Coordination Unit (SP1)
- Policy Framework (SP2)
- Licensing Models (SP2)
- Modern Data Capture Methods (SP5)
- Data Storage Processes (SP5)
- Metadata Standards – ISO 19115 (SP6)
- Establishing Partnerships (SP7)
- Identify Key Stakeholders (SP9)

Outcomes

- Increased Range and Scope of Authoritative Data
- A Critical Mass of Centrally Coordinated Data
- Cost Reduction Through Productivity Improvements
- Ability to Monitor and Measure Progress Towards Achieving the SDGs

Figure 4.1: *The overall structure for the Data Strategic Pathway - showing the four key elements, guiding principles, actions and interrelated actions, and the tools provided in the Appendices to support and achieve the outcomes.*

4.1 Introduction

With so much geospatial data becoming increasingly available, it is important to have guidelines and a framework to manage it consistently.

Geospatial information[2] reflects the physical world in which all human, economic and environmental activity takes place. It provides the digital version of our world - without which a digital economy is not possible. Geospatial information describes the physical location of geographic features and their relationship to other features and associated information such as census data. Geospatial information is presented in many forms and mediums including maps, satellite imagery and aerial photography. It describes the connection between a place, its people and their activities. It illustrates what is happening - where, how and why, and can be used to examine what has happened in the past, and to create likely future scenarios.

Geospatial information is collected and managed using a wide range of enabling technologies and tools, such as Geographic Information Systems (GIS), databases, photogrammetric software, satellites, mobile devices, and other ground-based, ocean-based and airborne sensors, etc. GIS in particular, is typically used to manage geographic features in 2 and 3-dimensional space. It is also used to visualize the dynamics of the environment as a series of data of the same area captured over time (4-dimensional space).

In many countries, there is a steady increase in the number of government departments, businesses and academic organizations that collect geospatial information, as well as end users that are continually adding value to data and redistributing it as new products. Decision-makers are increasingly recognizing the opportunities afforded through the ability to harness this information for evidenced-based policy setting, and integrated and transparent government services.

Most organizations use geospatial information in the course of their activities. However, when geospatial information is managed independently and not interoperable with other agency data, it is difficult to effectively bring together geospatial initiatives from the many agencies into collective action plans. This strategic pathway is a guide for data custodians (and producers) to manage their geospatial information using a whole-of-data lifecycle management approach (e.g. to organize, plan, acquire, harvest, curate, catalogue, analyze, integrate, aggregate, publish and archive this information) so that it can be easily incorporated, discovered and used/reused for decision-making.

This objective is crucial. Every decision made, every event or activity in our daily lives, occurs at, or as a result of, a geographic location. Whether it is to determine the best site for a new hospital, choose the location of a new business venture, stage a community event or respond to an emergency – geospatial information is inherent in decision-making.

For this reason, geospatial information needs to be integrated across government processes and services, and managed according to guidelines within the broader data, governance and technology frameworks of government. Once integrated, geospatial information becomes a powerful tool for

[2] For this strategic pathway, the terms 'geospatial data' and 'geospatial information' are used interchangeably in general contexts. In specific contexts, 'Geospatial Data' refers to unprocessed facts and figures; 'Geospatial Information' refers to data that has been processed, organized, structured and presented in a meaningful way.

determining appropriate policy interventions and decisions for a wide variety of government services that contribute to economic growth, national security, sustainable social development, environmental sustainability and national prosperity.

4.2 Context and Rationale

Geospatial information has historically been managed in silos, making data integration and exchange problematic. Cross-sector cooperation, standards and multidisciplinary collaboration can overcome these issues.

Authoritative geospatial information is often managed in closed systems of government, resulting in a multitude of information silos. Because of this, geospatial information is often not discoverable nor interoperable, making data integration and exchange problematic. Over time, duplicated information can arise resulting in inconsistent data versions, gaps in data coverage, a waste of resources, and an inability to effectively and efficiently integrate geospatial information across a broad range of services. Changes to reference datum can pose a challenge in the data supply chain. Sometimes the information required is not collected at all. There needs to be more institutional collaboration, interoperability and integration across the various national data information systems and platforms that exist. This includes a broader sharing of ground station and satellite data, data vocabularies, metadata and the adoption of open-source formatted data to promote data reuse.

Cross-sector coordination, multidisciplinary collaboration, and standards are required to overcome problems associated with data integration and overly complicated data supply chains, particularly when organizing, planning, acquiring, cataloguing, aggregating, integrating, curating, analyzing, publishing and archiving geospatial information.

Multidisciplinary issues like climate change, disaster risk reduction, emergency response and urban and regional planning require a coordinated effort across multiple government departments. Activities will often overlap when dealing with complex planning systems, such as implementing and operating an integrated flood risk management system or improving transportation infrastructure.

While organizations use geospatial information in the course of their activities, this information is typically managed independently and not interoperable with other agency data. This means it is not possible to effectively bring together geospatial initiatives from the many agencies into collective action plans.

Moreover, planning and risk management information will often be paper based. In addition to being difficult to share, paper maps are costly to create, time consuming to update and risk becoming obsolete and thereby that they may adversely impact decision-making. Paper-based information digitization should be encouraged and preserved as its conversion to digital formats would allow future updates, re-use and re-purpose.

4.3 Approach

The approach to effective data management is through well-defined data supply chains and custodianship guidelines for organizing, planning, acquiring, integrating, curating, publishing and archiving geospatial information.

With so much geospatial data becoming increasingly available there is an essential requirement to deliver a well-coordinated program for the collection, management and distribution of geospatial information. Being able to leverage geospatial information effectively and efficiently is what will afford and sustain increased economic activity and growth over the longer term.

Considerations towards coordinating the acquisition of geospatial data, and management of geospatial information, are an essential part of a country's geospatial strategy (See SP1: Action 1.6.6). These considerations are not only about what data to collect, but also about how to manage and maintain the data throughout its lifecycle (See Section 4.6.9).

In this strategic pathway, the approach for enabling data custodians to meet their data management, sharing and reuse obligations is through the execution of well-defined data supply chains and custodianship guidelines for organizing, planning, acquiring, aggregating, integrating, curating, publishing and archiving geospatial information (Figure 4.2).

The approach includes four key elements that are a guide for countries to achieve effective and efficient cross-sector and multidisciplinary geospatial data management, sharing and reuse. These elements include the implementation of priority fundamental data themes, guidelines for custodianship, acquisition and management, streamlined data supply chains and well-coordinated data curation and delivery mechanisms. These elements are explained in more detail in Section 4.4 below.

The approach includes strategic pathway actions that are recommended as a means to achieve the four key elements. The actions, which are underpinned by guiding principles, provide the step-by-step guidance to implement and achieve the desired outcomes. While most of these actions may be unique to this strategic pathway, there are several interrelated and/or prerequisite actions detailed in other strategic pathways that may also need to be completed. Tools to assist in completing the actions are available in the appendices to the strategic pathway. The approach for Strategic Pathway 4: Data is illustrated in Figure 4.2 and explained in the following Sections.

The actual implementation approach of each strategic pathway action will depend on country-specific needs, which may be influenced by country priorities, existing capabilities, resourcing potential, culture and other practicalities. Whatever the implementation approach, each action should reference the guiding principles below (See Section 4.5) as these describe what is important for effective and efficient geospatial information management.

Outcomes

- Increased range and scope of authoritative geospatial information available for decision-making and policy setting
- A critical mass of coordinated data to support national development and innovation
- Cost reduction through productivity improvement
- Ability to monitor and measure progress towards broad socio-economic benefits including the SDGs

Elements

- Data Themes
- Custodianship, Acquisition and Management
- Data Supply Chains
- Data Curation and Delivery

Tools

- Global Fundamental Geospatial Data Themes
- Data Theme Description Template
- Data Inventory Questionnaire
- Dataset Profile Template
- Gap Analysis Matrix
- Data Theme Road Map Template
- Principles for Data Custodianship Policy
- Data Governance Roles and Responsibilities
- Elements of a Data Management Plan
- Metadata Creation Checklist
- Data Release Guidelines
- Guidance for Improving Geodetic Infrastructure
- Global Statistical Geospatial Framework
- Guidance for Geospatial and Statistical Integration

Guiding Principles

- Data Governance
- Consistent Identification
- Data Quality Management
- Metadata
- Standards
- Accessibility
- Reusable Formats
- Authoritative
- Timeliness
- Provenance
- Integrity
- Demand Driven
- Efficiency
- Data security
- Respected Rights

Actions

Getting Organized

- Data Framework
- Data Inventory
- Dataset Profiles

Planning for the Future

- Data Gap Analysis
- Data Theme Road Map

Capturing and Acquiring Data

- Data Capture
- Data Acquisition Program

Managing Data Sustainably

- Data Custodianship Policy and Guidelines
- Data Governance
- Data Management Plan
- Maintained Metadata
- Data Release
- Data Storage and Retrieval Systems

Maintaining Accurate Positioning

- Maintained Geodetic Infrastructure

Integrating Data

- Geospatial and Statistical Integration
- Geocoding and Aggregation
- Data Supply Chains
- Data Interoperability

Interrelated Actions

- Geospatial Strategy (SP1)
- Country-level Action Plan (SP1)
- Governance Model (SP1)
- Geospatial Information Coordination Unit (SP1)
- Policy Framework (SP2)
- Licensing Models (SP2)
- Modern Data Capture Methods (SP5)
- Data Storage Processes (SP5)
- Metadata Standards (SP6)
- Establishing Partnerships (SP7)
- Identify Key Stakeholders (SP9)

Figure 4.2: *The approach to data.*

4.4 Elements

4.4.1 Data Themes

End users have a recurring need for fundamental geospatial data themes, as well as application-specific and socio-economic themes.

National priority and fundamental geospatial data themes, and the datasets that fall within each of these themes, are relevant to a broad range of applications; hence, end-users have a recurring need for this information. This need was recognized by the Member States of UN-GGIM as a key issue in 2012. As a result, UN-GGIM, from the efforts of the Working Group on Global Fundamental Geospatial Data Themes, has developed and adopted the Global Fundamental Geospatial Data Themes[3], a set of fourteen data themes that are the foundation to support global geospatial information management. Determined through global consultation and consensus, countries are encouraged to use and adopt these fourteen data themes, which can be adapted to align with national strategic and statutory needs and mandates.

The Fourteen Global Fundamental Geospatial Data Themes

In addition, data themes can include application data themes that are captured for a specific purpose, such as health and utilities; and socio-economic themes that are used for demographic studies.

4.4.2 Custodianship, Acquisition and Management

Data custodianship mandates responsibilities for the acquisition, management, maintenance and quality of the information.

Data custodianship is usually assigned to an organization, department or agency. It mandates certain rights and responsibilities for the collection of geospatial information and the management of this information on behalf of the community. The rights and responsibilities may include the right to set conditions for data release and responsibilities for the acquisition, management, maintenance and quality of the information. This is important, as geospatial information is only useful if the end-user is assured that the data is correct and up-to-date, or that it is the best available at the time. Data

[3] The Global Fundamental Geospatial Data Themes, adopted by Member States at the Seventh Session of the United Nations Committee of Experts for Global Geospatial Information Management, can be accessed at: http://ggim.un.org/meetings/GGIM-committee/9th-Session/documents/Fundamental_Data_Publication.pdf

custodianship usually includes the need to provide a recognized contact point for the distribution, transfer and sharing of the information. Mandating custodianship is one of the mechanisms used to avoid cross-government duplication in the acquisition and management of information. Custodianship is typically undertaken by subject-matter experts and the data in their care is recognized as the authoritative source.

Depending on national context, considerations should also be undertaken to harmonize and aggregate sub-national datasets such as when local organizations or agencies already have an established custodianship; or to establish a volunteering or crowdsourcing mechanism for data gathering particularly when data is non-existent or scarce.

4.4.3 Data Supply Chains

Data supply chains and their interlinkages need to be formalized and streamlined to improve the quality of information for end users.

Data supply chains refer to the flow of geospatial information from one organization to another. In a supply chain, organizations are referred to as supply chain nodes, and information flows are referred to as supply chain links. Each organization (node) will typically add value, such as updates, to the data before transferring the information on to the next organization (node). The types of interlinkages common to geospatial information flows are those between:

One level of government and another, such as between national and district levels;

- Organizations at the same level of government, such as between the geospatial organization and statistical organization;
- Departments, agencies, and volunteer geographic information providers and various projects; and between
- Departments, research institutes and private companies.

Many data supply chains that involve multiple organizations and different levels of government are not well-integrated and information becomes inconsistent and out-of-date very quickly. Consequently, information is not reliable. The resynchronization of datasets is a highly manual task, and therefore having formalized and streamlined, (i.e. standard data exchange, or ETL, Extract Transfer Load) processes in place from the outset, and data from an identified authoritative source, will save considerable effort later, minimize the supply chain length, and improve the overall quality and timeliness of information for users.

4.4.4 Data Curation and Delivery

Data curators typically aggregate data sources into a unified dataset ready for analysis.

Data curation and delivery refers to the process of enhancing and maintaining the value of data and delivering it to end users in a way it can be visualized and used. The main purpose of data curation is to ensure that data is continually updated, retrievable and functional for future purposes, reprocessing, or reuse. Data curation is usually assigned to data curators (or aggregators) who are responsible for collecting data from many different sources and then aggregating and integrating the data into an information resource, such as map portal. Curators then deliver shared data and information resources through data hubs for broad usage among stakeholders (partners, all sectors, and the public). Data curation has become increasingly important for data analytics, and the private sector cis often engaged in the curation of data to transform independently created government department data sources into unified datasets ready for analysis. Curated geospatial information is therefore often more valuable for decision-making than the individual geospatial datasets. Data delivery can be done using traditional approach through file system-based methods (such as coverages, or shapefiles) and direct connections to relational enterprise databases (such as Microsoft SQL Server, PostgreSQL, and IBM Db2) while modern approach use the internet and system approach to provide access to machine-readable data via Application Programming Interface (API, in formats such as JSON or XML) (see SP5 5.6.16).

4.5 Guiding Principles

By applying these guiding principles, data custodians will be able to manage and share reusable geospatial data, and in doing so, meet their obligations to government and the user community.

There are specific guiding principles and elements for managing geospatial information to promote consistent data governance, management, discovery, sharing and reuse so that data organizations may meet their obligations to government and the user community. These principles also need to be embedded into legislation, policies and directives for update/compliance and integration into business practices. The guiding principles for data are:

- **Data Governance:** High quality authoritative datasets are delivered through designated data governance roles and responsibilities. Data governance roles are mandated for each dataset to ensure responsibility for the integrity and quality of data.
- **Consistent Identification:** A common data dictionary, vocabulary, ontology and persistent identifiers are applied to the identification of data to enhance accessibility, manage effective use of data, and avoid duplicated collection or purchase.
- **Data Quality Management:** Data Quality Management processes are used to manage the currency, completeness, accuracy and consistency of data for a specified purpose.
- **Metadata:** Appropriate metadata is used to describe and discover geospatial data, including content, geographic extent, purpose, characteristics, temporal reference (currency and/or time) and provenance etc., together with contact details for further information.
- **Standards:** Appropriate and recognized standards are adopted and enforced throughout the data lifecycle to enhance integration and interoperability of individual and disparate data sets.
- **Accessibility:** Easy, efficient and equitable access to spatial data through common geospatial platforms and portals where technology, data formats, organizational arrangements, licensing, location, costs and conditions do not inhibit its use.
- **Reusable Formats:** Data is in a form suitable for further value-adding by internal and external users.

- **Authoritative:** Data is managed responsibly and efficiently by the designated data custodian to eliminate the proliferation of duplicate data sets. The notion is to collect once and use many times.
- **Timeliness:** Data is available for use in a timely manner as required by the application, including 'just-in-time'.
- **Provenance:** The origin and quality of data is readily accessible to the user, facilitating both human and machine-to-machine readability. via metadata so that they can determine if it is 'Fit for Purpose'
- **Integrity:** The interrelationships between data themes collected by multiple agencies are managed with topological integrity and with consideration to currency.
- **Demand Driven:** Data acquisition and maintenance is aligned to user needs and requirements to achieve optimal resource allocation.
- **Efficiency:** Geospatial data products are differentiated as close as possible to the user to create more opportunities for reuse along the supply chain, and appropriately tracked by versioning.
- **Data security:** Data are held with adequate provision for long-term care including disaster recovery and backup procedures, are disposed or archived in accordance with government regulations, and considering technological advancements.
- **Respected Rights:** Confidentiality, privacy, intellectual property rights and the security of sensitive information are protected and or anonymized prior to sharing, and the sharing of Indigenous knowledge is contingent upon consent of the knowledge holders in accordance with the principles of justice, democracy, respect for human rights, equality, non-discrimination, good governance and good faith.

4.6 Actions

The strategic pathway actions are recommended as a means to achieve the four key elements of the Data Strategic Pathway.

The strategic pathway actions are recommended as a means to achieve the four key elements of data. They are a guide to best practice collection and management of integrated geospatial information. Country-specific needs may be influenced by factors such as country priorities, existing capabilities, resources, culture and other practicalities. These will influence approaches for implementing each strategic pathway and their related actions.

For ease of use, particularly to assist countries in the initial and early stages of developing and strengthening their national geospatial information management arrangements, the actions are presented in a sequential step-by-step structure. A road map illustrating this order and where the actions typically occur and are completed, is presented in Figure 4.3. However, it is acknowledged that countries, depending on existing national arrangements, may also wish to start their actions at different steps along the pathway, and in a different sequence. Therefore, a less structured road map is additionally presented in Figure 4.4.

Some actions may have interrelated and/or prerequisite actions that need to be achieved prior to, or in conjunction with, the strategic pathway actions. These interrelated actions are also illustrated in Figures 4.3 and 4.4, are referenced in the text, and detailed under other strategic pathways.

Whatever the implementation approach, each action should take into account the guiding principles in Section 4.5, as these describe drivers for attaining effective and efficient geospatial information management.

The actions for the Data Pathway are divided into six categories, which are:

1. Getting Organized
2. Planning for the Future
3. Capturing and Acquiring Data
4. Managing Data Sustainably
5. Maintaining Accurate Positioning
6. Integrating Data

The following actions are typically used to address gaps in capability. They serve as a guide to building the necessary capacity to strengthen integrated geospatial information management processes and systems.

Figure 4.3: *Data includes several actions and tools designed to assist countries to achieve the best practice collection and management of integrated geospatial information. The actions are divided into six categories and reflect the order with which these actions are typically completed.*

Figure 4.4: *Data includes several actions and tools designed to assist countries to achieve political endorsement and strengthened institutional mandates for building a cooperative data sharing environment. The interrelated actions provide key linkages to other strategic pathway actions.*

1 Getting Organized

4.6.1 Data Framework

The Data Framework provides a way to organize geospatial information so that it can be accessed easily and meaningfully.

The Data Framework is a methodology for organizing a country's geospatial, statistical and other information so that it can be accessed and used. This is important. Being able to find information and understand its purpose is critical to good decision-making. There are typically three classification tiers to a Data Framework:

- Fundamental data themes, such as transportation, are required for a broad range of decision-making applications, and for which users have a recurring need. The 14 Global Fundamental Geospatial Data Themes provide a valuable and comprehensive starting point;
- Application data themes that are used for specific operations or activities, such as flood models, required for specific studies; and
- Socio-economic data themes that provide demographic information, such as census and population data.

The data themes are essentially categories for grouping geospatial datasets in each classification category. The themes make it easy for users to locate information and are important to structure and enable machine-readable data catalogues so that information can be located easily by search engines.

The Data Framework records Data Theme Descriptions in a way that makes it easy for users to understand what data is available within the theme, and the purpose for which it can be used. A Data Theme Description typically includes: (a) access, pricing and licensing categories; (b) the data custodian; (c) data characteristics, such as data structure, accuracy, and coverage; (d) data standards and regulations; and (e) the purpose for which the dataset is best used (see SP6: Action 6.6.9).

Additional guidance for defining data classes and models can be found in the Spatial Data Infrastructure (SDI) Manual for the Americas, provided by the Permanent Committee for Geospatial Data Infrastructure of the Americas (PC-IDEA, now UN-GGIM Americas) at the Tenth United Nations Regional Cartographic Conference for the Americas in New York in August 2013.[4]

Importantly, data themes, particularly fundamental data themes, should be considered when developing the nation's Geospatial Information Management Strategy (See SP1: Action 1.6.6) as they provide the foundation (base) information for achieving national goals and international commitments, such as the SDGs.

---

[4] E/CONF.103/14
https://unstats.un.org/unsd/geoinfo/RCC/docs/rcca10/E_Conf_103_14_PCIDEA_SDI%20Manual_ING_Final.pdf

In some situations, digitization of geospatial datasets may be required to collect data for each fundamental data theme. Nevertheless, new methods and emerging technologies can be adopted to fast-track the digitization process (See SP5: Action 5.6.6 to 5.6.8).

The Global Fundamental Geospatial Data Themes are available here: http://ggim.un.org/documents/Fundamental%20Data%20Publication.pdf

An example of a Data Theme Description Template for the Administrative Boundary Theme is provided in Appendix 4.1.

4.6.2 Data Inventory

A Data Inventory enables governments to fully understand the extent of their data holdings.

Conducting an inventory of all geospatial data and information held by institutions provides a valuable means for a country to fully understand the extent of its national data holdings. The Data Inventory should record, as a minimum, the metadata and spatial data format, currency, accuracy, ownership, extent/coverage, coordinate reference system and the purpose for which the data is primarily used. This information can then be incorporated into the Data Framework (See section 4.6.1).

It is important to differentiate: (a) what datasets are used by organizations as a reference only, as this highlights the benefits of data sharing; and (b) what datasets are collected and managed by organizations, as this will identify if there any areas of duplication in data collection.

Not all government geospatial information will currently be in digital format. Some of the country's wealth of knowledge may well be stored as paper-based documents/maps. This information will make a valuable contribution to decision-making capabilities when it is converted into digital form. It should therefore be considered in the data inventory process.

An example of a Data Inventory Questionnaire is provided in Appendix 4.2.

4.6.3 Dataset Profiles

Dataset Profiles improve the accessibility and use of geospatial information.

Dataset Profiles are created from the information collected during the data inventory. Like Data Theme Descriptions, Dataset Profiles are often incorporated into a data catalogue so that end-users can determine if the dataset will suit their purpose. This improves the accessibility and use of geospatial information. Dataset Profiles include specific information about the content of the dataset and can include:

- **Description:** General overview of the dataset theme and content.
- **Dataset Uses:** Lists the general purposes for which the dataset is intended to be used.
- **Current Status:** Date created and last modified.
- **Standards and Specifications:** Data and metadata standards applied.
- **Access and Licensing:** Conditions under which the dataset can be used.
- **Data Theme:** The Data Framework theme to which the dataset belongs.

- **Update Strategy:** Provides an indication of currency and when updates will be made available.
- **Dataset Custodian:** Organization responsible for acquiring and managing the dataset.
- **Data Curator (Aggregator):** Organization responsible for data aggregation, integration.
- **Distributors:** Responsible for delivering datasets to end-users.

Note: It is plausible that the data custodian, curator and distributor is the same organization.

An example of a Dataset Profile Template is provided in Appendix 4.3.
Note: Dataset Profiles are in human-readable formats; whereas metadata is in machine-readable formats.

See Interrelated Action on Metadata Standards - ISO 19115 (SP6).

4.6.4 Data Gap Analysis

A Data Gap Analysis is used to identify the strategies required to address the gaps in geospatial information capability.

The Data Inventory describes what datasets a country has, but not the actual data needs of the users, nor how to achieve the required data outcomes. A Data Gap Analysis is required to organize information so that it is simpler to identify the strategies required to address the gaps in geospatial information capability. Data gaps may include, incomplete data coverage in priority areas, inadequate accuracy, inconsistent geodetic referencing system/projection, limited integration between land parcels and addresses, restricted access to proprietary data, for example, utilities data, difficulties maintaining adequate levels of currency for building footprints, missing data themes, etc.

A Data Gap Analysis considers:

- **Current Situation:** This information is generally derived from conducting a SWOT and PEST analysis in a workshop situation with a broad stakeholder group. In this way, the current situation is viewed from multiple perspectives (see SP9: Action 9.6.1).
- **Desired Future State (Strategic Goals):** These are the goals a nation is endeavoring to achieve. Goals are generally documented in the Geospatial Information Management Strategy (See SP1: Action 1.6.7) and are associated with specific objectives that are aligned to the strategic priorities of government. In addition, the strategic goals consider stakeholder and end user needs, determined through stakeholder identification, analysis and surveys (See SP9).
- **Gaps in Capability (Challenge)**: These are the challenges that need to be overcome in order to achieve the goals and meet stakeholder and end user needs. Gaps may include lack of a required dataset, inconsistent data models, inability to share data effectively, etc.

- **List of Actionable Strategies:** A list of strategies, such as the collection of new data or the implementation of a new geodetic datum, which are required to move countries from their current status to a desired future state where geospatial data resources are concerned. These strategies provide input into the Data Theme Road Map (See SP4: Action 4.6.5) and are often included as actions in the Country-level Action Plan for strengthening geospatial information management (See SP1: Action 1.6.9).

An example of a Gap Analysis Matrix is provided in Appendix 4.4.

See Interrelated Actions on a Geospatial Strategy (SP1); Identify Key Stakeholders (SP9); Data Theme Road Map (SP4); and Country-level Action Plan (SP1).

4.6.5 Data Theme Road Map

A Data Theme Road Map is a powerful strategic tool for coordinating cross-government activities that will lead to strengthening integrated geospatial information management.

The preparation of a Data Theme Road Map is a powerful strategic tool for coordinating cross-government activities that will lead to strengthening integrated geospatial information management.

The Data Theme Road Map documents the major steps and milestones required to reach the desired outcomes and close the capacity gaps determined during the Gap Analysis (See SP4: Action 4.6.4). This high-level document serves as a communication tool to help articulate the strategic thinking behind the goals and the plan for getting there. As a minimum, the Road Map typically includes (Figure 4.5):

- A vision for what the datasets within the Data Theme will be in the future. The vision is usually driven by particular use cases i.e., what the data is required for as a matter of priority.
- Short term goals to be achieved and associated milestone dates.
- Activities to be undertaken, such as improving data quality, data supply chains, standards compliance, delivery mechanisms, policies associated with data release, retention and sharing, archive management for reuse and stakeholder engagement.
- Assignment of responsibility for change management:
- A Data Theme Sponsor with responsibility to engage with dataset custodians to guide initiatives to move each Data Theme towards the stated vision.
- An owner (usually the custodian) with responsibility to implement initiatives to improve the dataset.
- The outcomes associated with each action.
- Funding situation for each action and/or activity.

The Road Map is a key tool for achieving stakeholder buy-in. It should, therefore, achieve some immediately measurable benefits whilst progressing towards the eventual desired future state. Additionally, acquiring, enhancing and updating geospatial information is a costly and resource intensive process. The Road Map is typically designed to be completed over a three to five-year period and reviewed annually.

Figure 4.5: *Elements of a Data Theme Road Map*

Priorities will differ by country and should be based on the national political, social and economic drivers. Typically, the Global Fundamental Data Themes will be seen as the most important because they have an impact across a broad range of applications and stakeholders. They are also more likely to have a positive impact on economic growth. With benefits from initial efforts realized, it may be possible to add more data themes over time.

Within the Global Fundamental Data Themes, the imagery theme is likely to be a priority for many countries, as it is used to create other Data Themes, such as the digitization of buildings and settlements. While imagery is costly, there is a growing range of free and low-cost Earth observation satellite data that can be used by countries to initiate the transformation to digital geospatial information use and management (See SP:7 Action 7.6.2).

The data custodian will be integral to developing the Road Map for the data theme/dataset under their responsibility and therefore, these custodians should be determined at the outset.

An example of a Data Theme Road Map Template is provided in Appendix 4.5.

See Interrelated Action: Establishing Partnerships (SP7).

3 Capturing and Acquiring Data

4.6.6 Data Capture

Determining the best data capture methods to implement is an evolving practice. Each country will have a unique starting point - what is a suitable method for some, will be impractical for others.

Geospatial information is collected in various formats, typically - raster (images) and vector (lines, point and polygons). There are various methods for capturing and collecting new datasets, as well as enhancing and maintaining existing datasets. The chosen method will be dependent on the nature of the geographic feature, level of detail required, ease of updating, available resources and skills capacity, and budget.

Determining the best data capture methods to implement is an evolving practice. Each country will have a unique starting point - what is a suitable method for some, will be impractical for others. Data capture methods are also considered under Strategic Pathway 5: Innovation. Methods are discussed in terms of the potential to fast-track data collection and updating, as well as opportunities to leapfrog traditional manual methods and move straight to automated processes and new levels of location intelligence.

As a guide, the following should be considered when embarking on a data capture program.

- Does the data already exist elsewhere in a usable format? This may seem an obvious question; however, data can be difficult to find within closed systems of government and therefore, a thorough investigation is required.
- What information is necessary to meet internal organizational business needs and; what do end-users require and how will they use the information? This is crucial. For example, while an organization may require point locations of buildings, an end-user may need the actual building footprints. In this situation it would be logical to capture the building as polygons and derive the point locations from the building footprints.
- Are there options to partner with another organization? See partnership options in Strategic Pathway 8 that could lead to potential savings and resource sharing.
- Who will be the Data Custodian and how will the data be managed in the long term? This is an important consideration, particularly for projects which often result in isolated geospatial investments because consideration was not given to the longevity and potential future value of the data.
- Can existing datasets from sub-national authoritative sources be harmonized to form a nation-wide dataset? Often similar or duplicate datasets exist withing a country, and often at different levels of government. To achieve comprehensive nation-wide data coverage, independent datasets such as addresses, land parcels, building footprints and administrative boundaries, can be aggregated to form a national dataset.

- Can information be collected using volunteer geographic information (VGI) or crowdsourcing methods, or generated through social media, mobile GPS? In some cases when there is either scarce and low quality or no data available, the possibility of establishing a mechanism for people's engagement in data gathering should be considered (See SP5: Appendix 5.7).

See Interrelated Action: Modern Geospatial Data Capture Methods (SP5).

4.6.7 Data Acquisition Program

A nationally coordinated Data Acquisition Program supports the most strategic investments in geospatial data resources.

Each year government agencies purchase or collect a wide range of data and geospatial information products, such as aerial photography, satellite imagery, topographic mapping, and hydrographic surveys. The procurement of this information is often not coordinated or centralized, resulting in significant potential for the same information to be purchased, digitized and/or duplicated.

Geospatial information can be expensive, and given the potential for overlap and duplication, it is imperative that agencies adopt a more coordinated whole-of-government approach to ensure a more strategic investment and use of geospatial resources.

One mechanism to achieve this is to establish a National Data Acquisition Program, where organizations (including project teams) can register their requirements at agreed intervals. In this way the planning, recording and acquisition of geospatial information can be channeled through one governance conduit enabling clear oversight of data acquisitions nationally.

National-level acquisition may also be established through strategic data partnerships to combine data collected and managed at the sub-national level to create a seamless nation-wide visualization. In such cases, data management workflows are designed to enable federation and aggregation of dataset at the national level. The main objective of the Data Acquisition Program at different levels of government is to reduce costs associated with the capture, storage and management of geospatial information by:

- Procuring data once and in such a way that it can be used many times.
- Having clear oversight of nationally/internationally funded projects that acquire geospatial information.
- Acquiring multi-user and open licenses for data (such as imagery) to enable reuse by many organizations.
- Maintaining a single version of key government geospatial information.
- Leveraging greater value through economies of scale and consolidated spending.
- Establishing an annual program of work where organizations can register their requirements annually.

- Fostering collaboration where organizations have similar requirements, such as for multi-user licensing for imagery, in order to leverage greater value through economies of scale and consolidated spending.
- Engaging in Volunteer Geographic Information (VGI) and crowdsourcing data collection and data quality assessment initiatives that target a wide range of government programs.

A National Data Acquisition Program is managed with consideration for cross-government needs and in the national interest. The program may be administered by the Geospatial Information Coordination Unit (See SP1: Action 1.6.2) and will include representation and input from across key government stakeholders to decide on acquisition priorities and expenditure.

The Geospatial Information Coordination Unit will typically take responsibility for the business plan, management of finances and achievement of outcomes. Essentially, the Geospatial Information Coordination Unit can act as a community of practice for geospatial data providers and conduct the following tasks:

- Calls for geospatial data requests annually from organizations across the government sector.
- Maintain a register of all government geospatial data requests so as to avoid duplication and facilitate sharing and reuse.

The actual acquisition of geospatial data and imagery is typically undertaken by the survey or mapping department (or equivalent) that is responsible for preparing the acquisition plan based on stakeholder submissions (needs and priorities), designing procurement specifications, and validating the quality of the information as it is received.

A new procurement/funding model may be required to administer the National Data Acquisition Program. Options include:

- Agencies fund and manage their own geospatial data capture program but register their data capture program with the National Data Acquisition Program to avoid potential duplication.
- The National Data Acquisition Program receives an allocation of funds from central government to procure data on behalf of all organizations.
- Organizations contribute funds to the annual program in proportion to their data requirements.
- Commercial revenues from value-added data are redirected back into the National Data Acquisition Program.

See Interrelated Action: Geospatial Information Coordinating Unit (SP1); and Establishing Partnerships (SP7).

There may also be opportunities to develop data partnerships and strategic alliances (See SP7: Action 7.6.1 and 7.6.2) and negotiate more effective and consistent data licensing and copyright conditions for imagery, which may be more cost effective and result in reducing barriers to sharing and re-use of data within government.

4.6.8 Data Custodianship Policy and Guidelines

A Custodianship Policy provides a means of accountability for the production and delivery of reliable geospatial information.

A Data Custodianship Policy and Guidelines provide advice on the principles of custodianship and the responsibilities of custodians for the management of information products. They can include responsibilities for end-users when acquiring and using information. Guidelines are generally based on experience and simply reflect good practice, convention and convenience.

A custodian has the responsibility for data which includes security and privacy. There are significant advantages to be gained through having a Data Custodianship Policy, as it provides a means of accountability for the production and delivery of reliable geospatial information. For government agencies, this means being recognized as the authoritative source of the information. For end-users this means increased confidence in the accuracy, completeness and currency of government information. The data Custodianship Policy and Guidelines should follow the F.A.I.R. principles of Findability, Accessibility, Interoperability, and Reusability for scientific data management and stewardship (Wilkinson et al, 2016).

A Data Custodianship Policy also eliminates unnecessary duplication when capturing and maintaining geospatial information, as the policy typically mandates responsibilities to a single agency, or to custodians working in partnership. This means funds previously spent on duplicated activities can be reallocated to other priorities.

An example of Principles for Data Custodianship Policy is provided in Appendix 4.6.

4.6.9 Data Governance

Data Governance results in improved productivity and efficiency for an organization.

Organizations responsible for geospatial information management often have limited budgets and resources to regularly collect and update the diverse geospatial information in their care. Implementing Data Governance can result in improved productivity and efficiency for an organization. Data Governance[5] is a logical structure for classifying, organizing and communicating activities associated with making decisions and taking action on the management of geospatial information.

This model for Data Governance is different to the Governance Model described in SP1: Action 1.6.4. The Governance Model is designed to institute cooperation, visibility and alignment of decisions about geospatial information that has an impact across the broader government sector; whereas the Data

[5] For more information refer to the Data Governance Framework from the Data Governance Institute (2013) at: http://www.datagovernance.com/the-dgi-framework/

Governance model is implemented at the individual organizational level to govern how individual datasets are managed.

Most organizations will normally have some form of Data Governance – informal or formal. Organizations that formalize Data Governance will often do so when their data volumes are increasing, and being acquired from more diverse sources, creating an environment for data inconsistencies that need to be identified and corrected.

A Data Governance Framework considers the people, processes and technologies that are needed to manage, protect, enhance and deliver reliable geospatial information. Data Governance typically includes an organization's governance model and governing body, data governance roles and responsibilities, a set of procedures for data governance and a data management plan for executing those procedures (Figure 4.6).

Figure 4.6: *A Data Governance Framework*

Within organizations, Data Governance roles typically include an executive sponsor, IT steward, Data Steward, Data Custodian, Subject Matter Expert (sometimes referred to as the Data Owner), and the data producer.

An example of Data Governance Roles and Responsibilities is provided in Appendix 4.7.

See Interrelated Actions on Governance Model (SP1)

Understanding data lifecycle management is crucial to good Data Governance. Governing data at the point of creation and throughout the data lifecycle is one of the most challenging data management activities. The data lifecycle constitutes the data value chain; which is different from the data supply chain (See Section 4.6.17). The data value chain refers to the processes that are undertaken within an organization to add value to data; whereas the data supply chain refers to the transfer of data from one organization to another.

The data lifecycle will be different for each data and product type. However, there are stages common to all data and product types (Figure 4.7). These stages are associated with common overarching Data Governance roles (Appendix 4.7), and include:

- **Plan, Design and Specify:** This stage involves considering the end user and modalities for discovering the data (e.g., metadata or standards, etc.), and information to be accessed as well as planning data collection through to storage and publication methods. It includes the design of data models, and technical specifications and standards for capture and acquisition that are aligned with end user needs.
- **Create or Acquire Data:** Is the actual data creation and acquisition phase and can include general data maintenance and updating.
- **Validate and Verify:** Is the evaluation of data created according to documented guidance, policies, specifications, standards and/or legal requirements. Validation is the process of checking whether data or data models meet the required specification; and verification is the process for checking the accuracy of data entered into a dataset/system to test that it exactly matches the source information.
- **Reappraise:** Data that fails validation and verification is returned for correction. For example, data may be returned to the contractor if the data does not adhere to standards.
- **Archive:** Data is transferred onto a computer hard drive or shared network drive or storage repository for long-term retention. The archive typically contains data that is not actively used and there is a need to implement data security measures to protect data from unauthorized access, data corruption and accidental data destruction, modification and disclosure.
- **Dispose:** While information is typically archived for later use, some information has a finite purpose and can be discarded at a later date once it has served its purpose or is no longer valuable to the organization, after a mandated 'records retention' date.
- **Preserve:** Actions are undertaken to ensure long-term preservation and retention of the authoritative nature of data. Preservation actions should ensure that data remains authentic, reliable and usable while maintaining its integrity. Actions include data cleaning, validation, assigning preservation metadata, assigning representation information and ensuring acceptable data structures or file formats.
- **Migrate:** Migrate data to a different format or anonymize the data. This may be done to accord with the storage environment or to ensure the data's immunity from hardware or software obsolescence.
- **Store:** Data is stored in a secure manner where it can be routinely accessed for reuse. Data stores are typically publicly accessible data warehouses and there is a need to implement cyber security measures against.
- **Access, Use and Reuse:** Data is accessed by users on a day-to-day basis. This may be via internal systems or in the form of publicly available published information. Data is typically accompanied by a license agreement.

- **Analyze and/or Transform:** End users create new data from the original, for example by migrating data into a different format, integrating data with other data, or by creating a subset, by selection or query, to create newly derived results, perhaps for publication as information products.
- **Publish:** Data is published as traditional data products and value-added information products, in print form, as web services, Application Programming Interfaces (APIs), consumer and business Applications (Apps) etc. Data is typically subject to copyright laws.

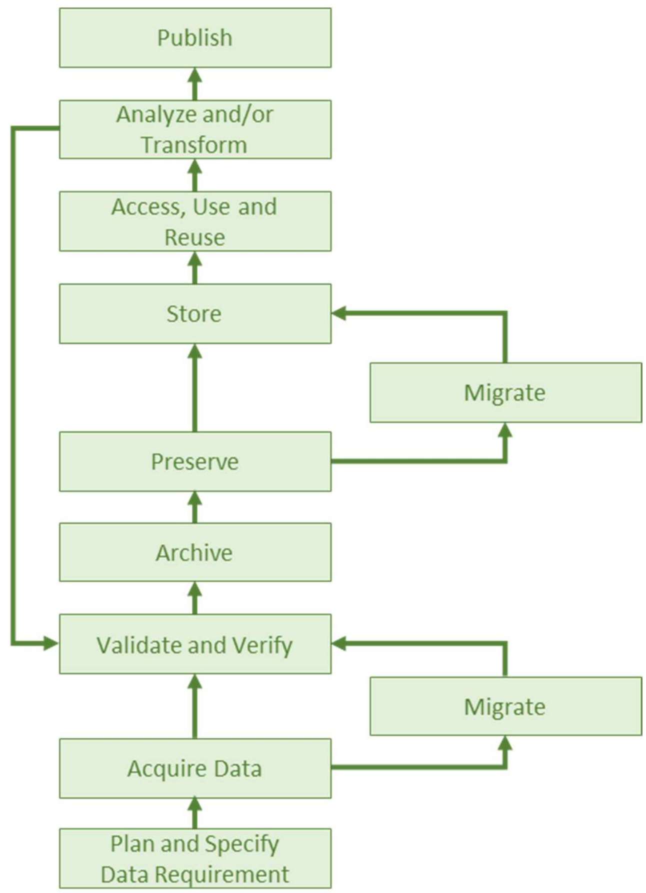


Figure 4.7: *Common Data Lifecycle Model*

4.6.10 Data Management Plan

Data Management Plans promote consistency in geospatial data management practices within agencies and across government.

Data management plans promote consistency in geospatial data management practices within agencies and across government. They are a strategic tool used to provide oversight for data assets as well as other ICT (Information and Communications Technology) assets so that maximum value is derived through appropriate governance, maintenance and protections.

Data management plans are a useful tool for overcoming organizational data silos. The guidelines in the plan facilitate data discoverability, accessibility and interoperability through uniform standards, definition of data models, processes and methods (See SP6: Action 6.6.9). They streamline processes through the implementation of standards and best practices, resulting in geospatial information that is suitable for frontline service delivery, analytics and decision-making.

Elements of a Data Management Plan are provided in Appendix 4.8.

The major elements considered in the Data Management Plan are custodianship, knowledge management, security, quality and accessibility (Figure 4.8).

Figure 4.8: *Elements of a Data Management Plan*

4.6.11 Maintained Metadata

Metadata makes it possible to access geospatial data and determine its usefulness for mapping and analysis.

Metadata is ‘data that describes other data’. It describes the origin of geographic data and tracks the changes to a dataset over time. Metadata refers to the ‘who’, ‘what’, ‘why’, ‘where’, ‘when’ and ‘how’ about individual geospatial datasets and how they have been collected (Nebert, 2004). Without metadata it is difficult to access geospatial data and, when located, impossible to determine its usefulness for mapping and analysis. Without proper metadata that describes critical characteristics of the data, such as the data capture, methods, quality, source, accuracy or temporal validity, and distribution, else the data might be unusable for users.

The concept of metadata can be likened to a library catalogue which includes metadata records that help a person discover, use and manage a collection of books, documents and other information. Metadata for digital geospatial data is no different; except that the emphasis is on the geospatial component and includes characteristics such as producer, extent, accuracy, datum, etc.

Metadata is managed at different levels including discovery, exploration and exploitation (Geospatial World, 2010; van Oosterom and Ziatanova; 2008):

- Discovery metadata is used to describe what data sets are available and provides a mechanism by which an agency can publish their data holdings.
- Exploration metadata allows users to explore data so that they can determine if the currency and accuracy of data are sufficient for their needs.
- Exploitation metadata includes those properties required to access, transfer, load, interpret and apply data in end-use applications where it can be exploited by the user.

The minimum amount of metadata recommended is the information contained in the Dataset Profiles (see 4.6.3 above). The form of metadata creation and its maintenance will depend on the size of an organization's data holdings and the patterns of data management within the organization. Traditionally, organizations with modest data holdings store metadata in conventional discrete word processing documents. This methodology meets minimum requirements. However, for large data holdings consideration should be given to employing machine-readable formats using more advanced aspects of GIS to extract aspects of the metadata from the data itself. Metadata should be created with a view to supporting the data lifecycle management and implementation format within a database or software system; an export (encoding) format for transfer between computers and; presentation formats for viewing. In this way, the system will support business and operational requirements, standard encoding for data exchange, and permit a number of 'report' views of the metadata that will satisfy the needs and experience of the different users.

Consideration should be given to using international metadata standards that use a structured exchange format and flexible presentation and software support for encoding (see SP6: Standards).

Metadata validation is an important business process. However, there are few tools available to automatically verify the accuracy of the metadata other than those tools for checking its syntactical structure. Human review is required and should be incorporated into an organization's business process.

The management of the metadata is best left with the data source, the agency that collects and maintains the data. However, metadata catalogues may be more appropriately managed by a central coordinating team, which can ensure consistent metadata management across organizations and thus ensure a consistent means of accessing data via a national data catalogue.

Although vital to a dataset's provenance, there is still a risk that data producers will consider the creation of metadata for their datasets too hard or time consuming to do so. This is particularly the case for organizations that may have hundreds of legacy datasets. Organizational managers may need to commit additional resources to the task, and be conscious that time spent documenting each dataset is a small expense when compared to the months taken to produce the dataset in the first instance. It is worth noting that metadata may be harvested automatically from the datasets themselves and assigned in bulk (e.g., owner details).

A checklist for Creating Metadata is provided in Appendix 4.9.

Metadata based on international standards will have far reaching 'long term' benefits for both data producers and the geospatial information user community. Strategic Pathway 6: Standards provide more information regarding metadata standards.

4.6.12 Data Release

Data Release Guidelines provide custodians with guidance on what data can be lawfully released in the national interest.

Data custodians are accountable for the release of information, and this responsibility is often mandated under a Custodianship Policy. However, custodians often need additional guidance on what can be lawfully released in the national interest. This is usually achieved by assigning an access category to each dataset. Access categories are typically:

- **Open:** Open Access - Open access to geospatial data takes precedence over restricted access unless there are specific, compelling reasons to restrict access.
- **Restricted:** Approval of an authorized officer is required. Decisions on restricting access are based on privacy, commercial sensitivity, national security, environmental sensitivity, cultural sensitivity or legislative requirements.
- **Confidential:** Internal Use Only – Access to authorized users only, where access is necessary to the function or activity of the collector/user.
- **Not to be Released:** A default position for all types of data and information until assigned a category.

When assigning access levels, there needs to be a balance between making information openly available for public good and economic growth; with the need to restrict access for the protection of individual rights and national security.

There are times when certain classes of geospatial data need to be withheld from public access and use. For example, personal information, defense establishments, culturally sensitive sites, detailed bathymetry of harbor approaches, and locations of endangered species. There are also times when withholding data can degrade decision-making processes, such as emergency planning and response, and environmental management.

While it is recognized that some data cannot be made public because of its sensitivity, data should still form part of the Data Framework and be managed by nominated custodians. Authorized users can be given access for business purposes and mechanisms can be introduced to ensure that privacy, national security and other sensitivities are not compromised.

Importantly, the issue is not about whether geospatial data should be collected and made accessible, but rather what restrictions will be applied to its usage and how this should be decided. This can be achieved by specifying a License Type.

An example of Data Release Guidelines is provided in Appendix 4.10.

See Interrelated Actions on Licensing Models (SP2).

4.6.13 Data Storage and Retrieval Systems

Information management has become more difficult in recent times with larger volumes of data and limited ICT budgets for storage and retrieval.

In some developing countries, valuable geospatial data may be stored as hardcopy records of maps, plans, ledgers, etc. These hardcopy products are often the only records of land boundaries, urban development schemes and construction plans, etc. This paper-based information consumes a lot of physical storage space, is prone to damage and deterioration, and can be vulnerable to natural disasters.

Digitizing these documents has many benefits - reduced physical storage requirements, lower document management costs and timeliness of processes, increased document security and reduced risk of damage, better knowledge management and search capabilities, opportunity to build in regulatory compliance, and opportunities for teamwork through digital collaboration technologies.

However, moving to a digital environment brings its own risks and the need for secure data storage and retrieval processes. The digital storage capacity within government organizations of many developing countries is often under resourced, increasing the risk of data becoming lost. Some organizations store data on computer hard drives and are only accessible to the computer owner/user. More concerning is that this information sometimes represents the only copy, resulting in a high risk of information loss. In addition, due to their file size, digital imagery is often stored on transportable hard drives, while historical imagery may remain archived on film. Both of these situations present long-term risks due to disk corruption and the continual deterioration of film.

Information management has become more difficult in recent times. Governments typically need to retain larger volumes of data for longer periods of time to meet governance and compliance requirements, and they are faced with limited information technology budgets. Cloud storage provides much needed storage capacity for government. However, few organizations are connected, mainly due to a lack of awareness of the storage facility, and concerns regarding the security and sovereignty of data that is not kept on home soil. Cost may also be a limiting factor. The Future trends in geospatial information management: the five-to-ten-year vision report details that "the use of cloud computing has risen substantially in the last decade and has become the current favored storage option. Cloud computing offers many benefits (reduction in operating costs, reliability and scalability of service provision). However, there are also challenges, for instance quality and security issues are still to be solved or clarified before some kinds of data are transferred to the cloud" (UN-GGIM, 2015). For a discussion on the different types of data storage processes, solutions and innovative methods, see SP5: Innovation.

See Interrelated Actions on a Data Management Plan (SP4: Activity 4.6.10); and
Data Storage Processes (SP5: Activity 5.6.8).

When implementing storage and retrieval systems, the following guiding principles apply:

- **Security:** Data is maintained in a secure environment and transmitted through secure methods.
- **Maintain:** Data is held with adequate provision for long-term care including disaster recovery and backup procedures.

- **Disposal:** Data that is not required is disposed of or archived in accordance with Government record keeping legislation.
- **Retrieval:** Data can be easily retrieved in a form that is needed by authorized persons.
- **Compatible:** Information is stored in a way that is open formats so that it is compatible with multiple hardware and software systems that will be used to access it.

5 Maintaining Accurate Positioning

4.6.14 Maintained Geodetic Infrastructure

The geodetic infrastructure is a prerequisite for the accurate collection, integration and utilization of all other geospatial data.

The Global Geodetic Reference Frame (GGRF) provides the reference system for our planet. It is the foundation of virtually every aspect of the collection, management and use of national geospatial information and global monitoring of the Earth. The GGRF is a prerequisite for the accurate collection, integration and utilization of all geospatial data, as it allows the interrelationship of measurements taken anywhere on the Earth and in space. Geodetic infrastructure (also referred to as positioning infrastructure) allows users to precisely determine and express locations/coordinates on the Earth, and to measure changes to the Earth's surface over time.

To ensure the long-term sustainability of the GGRF, the United Nations General Assembly adopted resolution 69/266 on 'A global geodetic reference frame for sustainable development' in February 2015[6]. Since then, the UN-GGIM Sub-Committee on Geodesy (SCoG) has been established and has delivered a GGRF Roadmap[7] and Implementation Plan[8].

Without the existence of this predefined geodetic framework, the GGRF, almost every endeavor that depends upon positioning information would become daunting, from both a practical and financial viewpoint. For example, the GGRF supports all satellite positioning technology, and is underpinned by critical and unique globally distributed ground infrastructure including observatories and satellite tracking stations. The GGRF and related infrastructure underpins all geospatial data and provides the means for relating all cultural, administrative and geographic features to the one horizontal and vertical coordinate system (geodetic datum) that is used by a nation. Importantly this includes providing the positioning framework for the cadastral system.

At the global scale, the globally distributed geodetic infrastructure includes the network of high fidelity geodetic observing stations, many with multiple collocated observing techniques – Very Long Baseline Interferometry (VLBI), Satellite Laser Ranging (SLR), Doppler Orbitography and Radiopositioning

---

[6] Resolution on GGRF is *available at* http://ggim.un.org/documents/A_RES_69_266_E.pdf
[7] GGRF Road Map *available at* http://ggim.un.org/meetings/GGIM-committee/documents/GGIM6/E-C20-2016-4%20Global%20Geodetic%20Reference%20Frame%20Report.pdf
[8] GGRF Implementation Plan *available at* http://ggim.un.org/meetings/GGIM-committee/8th-Session/documents/Road-Map-Implementation-Plan.pdf

Integrated by Satellite (DORIS), Continuously Operating Reference Stations (CORS) and Gravimetry (Quantum, Superconducting and/or Absolute). These stations, their data archive and the data processing/analysis centers form the basis of the GGRF, and are owned and operated in a collaborative effort by many contributing organizations across the globe. Furthermore, they support precise positioning from Global Navigation Satellite Systems (GNSS) through the coordinate systems transmitted by the various GNSS.

At the national level, the geodetic infrastructure may include:

- Instruments such as regulations, standards and guidelines that, on one level, provide the authority and requirements for geodetic infrastructure practices and processes. At another level, they also determine and assign responsibilities for countries for the implementation and maintenance of national geodetic infrastructure.
- A geodetic datum that allows a point on the Earth's surface to be uniquely described or identified by its latitude and longitude. A geometric datum is based on a mathematical model of an ellipsoid and has reference points that are used to define the datum. A vertical datum will be based on a reference to sea-level or on a defined gravimetric potential value.
- A Geoid Infrastructure to maintain an accurate height system to allow economic and efficient height measurement by GPS in support of a range of applications in large scale construction projects, mining, agriculture, flood protection and food security. Gravity data is required to fully utilize the potential of GPS based Geodetic surveying.
- A practical realization of a geodetic datum enables access to the fundamental spatial reference system for all subsequent surveying requirements for cadastral, mining, mapping, engineering and related work.
- Geodetic data that includes reference point horizontal coordinates that support the positioning and mapping within a datum, height information, and the means of transforming different coordinate systems and relating different height systems, along with supporting metadata.

Geodetic services are provided by many countries which transform client data into precise position information in global or national reference frames with error estimates and other ancillary information.

Physical geodetic infrastructure, which includes:

- For some countries, components of the global geodetic infrastructure described earlier, including VLBI, SLR and DORIS (as they do need to be located in countries) as part of the geodetic infrastructure. These types of stations ensure the GGRF and its availability.
- Passive survey marks, of varying degrees of quality, ranging from low precision ground monuments to high precision GNSS pillars capable of measuring minute geodynamic motions.
- Active CORS - a permanently configured network of GNSS observing stations that support a range of positioning infrastructure purposes and applications.
- Information storage, maintenance, delivery and access systems including the transfer of data files, data downloads from Web portals, integrated web services and mobile applications.

- Suitably qualified and experienced personnel to operate and maintain the geodetic infrastructure.

Geodetic infrastructures differ between countries, often due to the nature of the geographic features (e.g., flat verses mountainous terrains) and economies, and because there are unique problems that need to be overcome to establish adequate geodetic infrastructure.

In developing countries, the high cost of establishing and maintaining geodetic infrastructure has meant that some countries have relied on temporary survey marks for project work. This has led to a lack of standardization and uniformity, and as a consequence, the geodetic infrastructure has become fragmented, unconnected to the national or regional framework and inadequate for accurate positioning.

In this situation, the question is 'What should countries do to arrive at an adequate geodetic infrastructure?' Appendix 4.11 provides background information and guidance for improving geodetic infrastructure.

> Guidance for Improving Geodetic Infrastructure is provided in Appendix 4.11.

6 Integrating Data

4.6.15 Geospatial and Statistical Integration

Integrated geospatial and statistical data enables good policy and decision-making.

Geography and statistics are implicitly interconnected in a conceptual sense. Demographics (population, education, health, labor), economics (trade, tourism, sales, salaries, etc.) and environmental information (ecosystems, climate, water quality, traffic, etc.) are all associated with a geographic location or region, e.g., municipality, postal district, country.

However, geography and statistics are frequently disconnected at logical and physical database implementation levels. This is partly because geospatial technologies are only used to a limited extent by some National Statistical Organizations (NSOs) and mostly for census-related work, but also because geospatial data and statistics are managed in separate organizational structures.

With the adoption of the 2030 Agenda for Sustainable Development, and subsequent implementation of the SDGs, there is growing appreciation for the value of integrating geospatial and statistical information for making good policy decisions that will lead to the achievement of the SDGs, and for monitoring and measuring progress towards their implementation.

There is also a growing expectation for governments to release location-enabled socio-economic data for broader business opportunities, innovation, and research and development. This is because geospatial and statistical data integration increases the level of understanding of the interrelationships between statistics, geographic phenomena and human endeavor through enhanced static and dynamic multi-dimensional visualizations. Such information also underpins the national Population and Housing Censuses undertaken by countries.

For these reasons, the United Nations Statistical Commission (UNSC) and UN-GGIM established the Expert Group on the Integration of Statistical and Geospatial Information (EG-ISGI) in 2013 and charged the Expert Group with the development of a Global Statistical Geospatial Framework (GSGF). Now completed and adopted by UN-GGIM and the UNSC, the GSGF facilitates the integration of statistical and geospatial information. It enables a range of data to be integrated from both statistical and geospatial communities and, through the application of its five Principles and supporting key elements, permits the production of harmonized and standardized geospatially enabled statistical data. The resulting data can then be integrated with statistical, geospatial, and other information to inform and facilitate data-driven and evidence-based decision making to support local, sub-national, national, regional, and global development priorities and agendas, such as the 2020 Round of Population and Housing Censuses and the 2030 Agenda for Sustainable Development.

The GSGF relies on critical inputs of fundamental geospatial data, and supplements these with other data sources as necessary. These data then serve to geospatially enable traditional and authoritative statistical data, and increasingly data from administrative and other sources, much of which comes from NSOs and administrative data custodians within the broader national statistical system.

The Global Statistical Geospatial Framework (GSGF) is available here:
http://ggim.un.org/meetings/GGIM-committee/9th-Session/documents/The_GSGF.pdf

Guidance for Geospatial and Statistical Integration is provided in Appendix 4.12.

The five Principles of the GSGF outline the broad processes by which a range of geospatial and statistical infrastructures and processes are applied to input data to enable integration. Firstly, the statistical data are geospatially-enabled to the finest level possible. Then, geospatial tools and methods, such as common geographies and common standards of good practice, are used to ensure the data are interoperable, accessible, and usable. The five Principles are:

- Use of fundamental geospatial infrastructure and geocoding;
- Geocoded unit record data in a data management environment;
- Common geographies for the dissemination of statistics;
- Statistical and geospatial interoperability; and
- Accessible and usable geospatially enabled statistics.

In addition to the five Principles, the GSGF features inputs, key elements and outputs that act as a bridge between statistical and geospatial professional domains, between National Statistical Organization (NSOs) and National Geospatial Information Agency (NGIAs), and between statistical and geospatial standards, methods, workflows, and tools (GSGF, 2019).

4.6.16 Geocoding and Aggregation

Geospatially enabling data – linking information to a location – is an important component of data integration.

Geospatially enabling data – linking information to a location – is an important component of data integration, as identified by Principle 2 of the GSGF. The linkage between a geographical point location or boundary and statistical records is achieved through a computational process referred to as geocoding and aggregation. Geocoding is used to link statistical data at a micro level, such as linking household data to a building, whereas aggregation is used at the macro level to represent a generalized interpretation of statistics by area, such as the demographics of a village, district, suburb or other administrative area.

Geocoding is used to geospatially locate statistical unit records to a high degree of accuracy. This is achieved by transforming a description of a location – such as a pair of coordinates, an address, or a name of a place – to a location on the Earth's surface. The land parcel or property is often used as the geographic feature with which statistics are associated, and the address and/or coordinate location of the parcel or property is typically used to link the two elements – statistics and location. Geocoding can be completed by entering one location description at a time but is mostly achieved by processing a table of records in bulk at one time.

There are several methods that belong to the geocoding family of transformations (Figure 4.9). The methodology is dependent on the type and format of the data, the geographic location geometry (point/polygon), and amount of information to be geocoded. Geocoding need not be an expensive task, and there are several commercial and open-source services available.

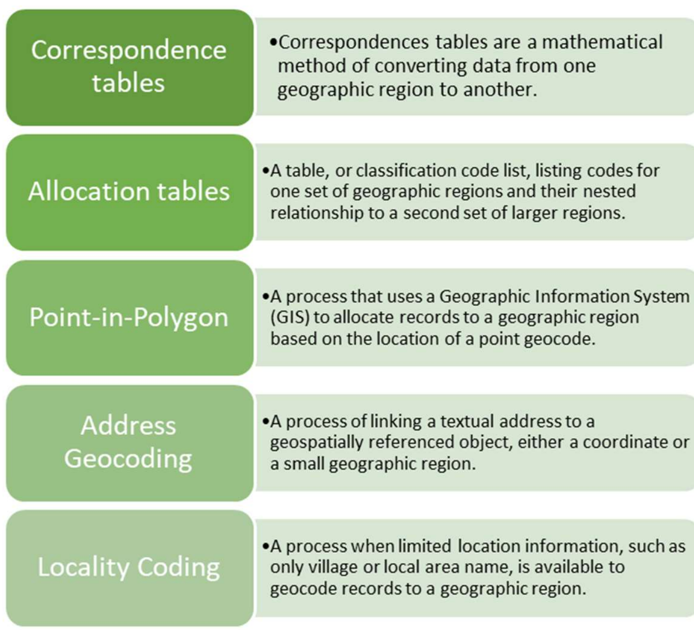


Figure 4.9: *Examples of geocoding methods available.*

4.6.17 Data Supply Chains

A Data Supply Chain Strategy is used to formalize data sharing, invoke supply chain partnerships, and establish service level agreements between organizations.

The exchange of information between multiple organizations participating in a national data supply chain is often not well documented. This leads to potential data duplication and repetition of human resources, software and storage overheads. To mitigate this situation, a national Data Supply Chain Strategy is used to formalize data supply chains. It encompasses four domains (Figure 4.10) (Arnold, 2016):

- **Business Domain:** The supply chain strategy delivers on the national business (or economic) outcomes required, such as return on investments and incentives for businesses. These are generally specified in the strategy along with the collective vision, mission and goals of the supply chain partners. The supply chain strategy also considers the value proposition to the end-user. This value stems from the collective efforts and activities of the supply chain partners.
- **User Domain:** The supply chain strategy considers the end-user requirements, the factors influencing product usage behavior, and the design criteria that will create the most value for the end-user, such as timeliness, content, coverage, semantics and accuracy.
- **Production Domain:** The supply chain strategy coordinates the external forces that impact on demand planning, data sourcing complexity, and the types of transformation processes required to make a data product. A compliance framework is required to support interoperability including data and technology standards, quality measures and metrics, and custodian roles and responsibilities. Collaboration with supply chain partners is a key component to sustaining production in the longer term.
- **Service Domain**: The supply chain strategy focuses on connecting people to products and services. It considers the integration of component products from multiple sources to create standard offerings as well as tailored solutions. A policy framework is required to manage open access to data products balanced with individual privacy, data sensitivity, copyright and intellectual property considerations (See SP2: Actions 2.6.7 to 2.6.12). These aspects are more complex in extended supply chain networks. Communicating product suitability may require a rating system that is meaningful in the end-user's context.

The national Data Supply Chain Strategy invokes supply chain partnerships and establishes service level agreements for updating and sharing data between organizations to ensure the ongoing currency of information. This is important. As soon as data is physically transferred from one agency to another, the data becomes stand-alone and either evolves into an obsolete dataset because it is not updated, or conversely, if it is updated the data becomes inconsistent in terms of currency with the parent data source. The Data Supply Chain Strategy formalizes consistent and timely data supply and exchange processes.

Geospatial data supply chains in many countries have evolved to become a complex network that is difficult to visualize and manage. Understanding the flow of geospatial data updates is crucial to maintaining, synchronizing and sustaining datasets over the longer term. However, creating supply chain models is not a simple task. Some datasets are stewarded from the bottom up, with local governments producing data that can be aggregated at the national level. Other datasets are managed through national programs with information transferred to the local level at specified intervals. In other cases, government agencies may acquire and use commercial data or collaborate with NGO's, citizens or volunteers.

Figure 4.10: The four domains to be considered when developing a National Data Supply Chain Strategy (adapted from Arnold, 2016).

Therefore, developing a supply chain strategy provides an opportunity to review current capabilities and make an assessment on areas that could be achieved more effectively and efficiently. It is also an opportunity to consider new digital trends, appraise new methods (see SP5: Innovation) such as open geocoding and normalization services to automate data integration (GIS Geography, 2020), and consider engaging with new supply chain partners, such as an external data aggregator or community volunteer geographic information providers. While simple Data Supply chain can be addressed using a simple Data-As-A-Service (DAAS) solution, most national context and data require more complex integration scenarios which are not easily resolved. There are three levels that join together to deliver a higher-level 'specifically' integrated supply chain strategy that is a plan for creating, delivering, governing and measuring geospatial data content, and is relevant to end-users at national, jurisdiction and business levels. The approach recognizes that different organizations have their own strategic drivers (operational mandates), as well as meeting data sharing directives in the national interest. The three strategic levels, as shown in Figure 4.11, are:

- A National Supply Chain Strategy that is collectively developed, managed, adopted and monitored by channel partners, such as the digital cadastre supplier/producer representatives from each jurisdiction. It is generally an aspirational strategy.
- A Jurisdiction (or sub-national or local government level) Supply Chain Strategy that encompasses national requirements as well as immediate jurisdiction requirements and formalizes local supply chain partner relationships.
- A Business Strategy (or Data Governance model – see Section 4.6.9) that is developed at an organizational level to deliver discretionary (or mandated) data products and services, as well as obligatory data products and services for jurisdiction and national end users.

The integrated strategic approach addresses the challenges of linking communications across supply chain channel participants of different orders (their level within the supply chain), and across broader policy initiatives and country-level development programs.

The Data Supply Chain Strategy is also an opportunity for organizations to evaluate their geospatial data holdings and strengthen their data management processes. Supply chain strategies often bring clarity to data custodianship and responsibility for specific upstream and downstream activities.

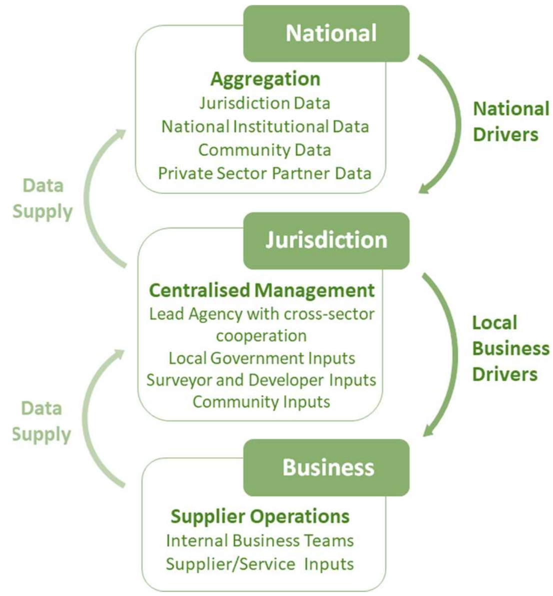


Figure 4.11: Integrated Supply Chain Strategy - three levels showing data flows and drivers.

Given that data duplication may have occurred in the past, the supply chain strategy additionally affords the opportunity to:

- consider which datasets represent an official dataset - fully endorsed by the government for specific purposes, such as for the purpose of reporting on the Sustainable Development Goals; or
- put in motion procedures to have datasets recognized as an official source.
- identify opportunities for benefits and economic return on investments.

The Data Supply Chain Strategy is also an opportunity for project sponsors to consider operating within a more sustainable supply chain framework rather than within an isolated project scope. Data supply chains are often a challenge for project teams because it is not always clear how their project-level data acquisitions, including community provided information, should be incorporated into mainstream data supply chains.

As government data supply chains can span several organizations, where to introduce project-level data can be problematic, particularly as each organization in the supply chain contributes to its value in some way. Adding to this complexity, information flows can be bi-directional across organizations at the same level of government as well as between organizations at different levels of government. Because different data vocabularies, standards and schemas can occur, exchanging and reusing project-level data is not necessarily straightforward.

There are several methods available to organizations and project teams to streamline supply chain linkages. The most commonly used include:

- Incremental Updating, where project data updates are forwarded to the data custodian for the inclusion/integration into the authoritative database.
- Database Versioning, where updates are performed by the project team on a copy of the custodian's database, which is then integrated into the authoritative source by the custodian.
- Direct Editing, where updates are performed through a defined process directly in the custodian's authoritative database.

With the exception of the direct editing approach, the integration of project-level data with the primary dataset is essentially a manual task. However, research is examining more automated methods to conflate project-level data acquisitions with authoritative sources (Cooperative Research Centre for Spatial Information, 2017). Current methods rely on the use of common standards so that data elements can be combined across datasets more readily.

In summary, the Data Supply Chain Strategy should:

- Assess existing supply chain capabilities relative to good practice.
- Consider the current and future business requirements of supply chain participants (contributors) as well as the needs of end-users.
- Evaluate new technologies and methods (See SP5: Action 5.6.3), and whether or not new core competencies are required (See SP8: Action 8.6.4).
- Consider new supply chain partnerships and community mapping programs and discontinue ineffective networks that lead to duplication of effort.
- Include a Supply Chain Road Map (multi-year) to improve supply chain interlinkages one step at a time. The Supply Chain Road Map should dovetail with the Data Theme Road Map.
- Gain buy-in and agreement for the Data Supply Chain Strategy from all participants.

Consider how quality will be managed and coordinated along the supply chain, including what standards need to be complied with (See SP6: Standards), which organization will be responsible for compliance, how quality control will be managed across the supply chain, and what procedures are required for quality assurance purposes.

4.6.18 Data Interoperability

Data interoperability is crucial to achieving integrated data supply chains.

As a key component of the F.A.I.R. (Findability, Accessibility, Interoperability, and Reusability) principles, data interoperability is crucial to achieving integrated data supply chains. Having data that is interoperable means that systems and services that create, exchange and consume data have clear, shared expectations for the contents, contexts and meaning of the data (Data Interoperability Standards Consortium, 2018). In addition to promoting standardization for data sharing and reuse, interoperable data supports multidisciplinary knowledge integration, discovery, innovation and productivity improvements. Data interoperability is essential throughout the data life cycle, from data capture, and publication, to ensure it can be readily available, integrated and used in a variety of context of decision- and policy-making purposes including early warning, crisis management, or long-term planning.

To be interoperable the data will need to use community agreed formats, language and vocabularies. The metadata will also need to use agreed standards and vocabularies and contain links to related information using identifiers (ANDS, n.d.). For more information on common vocabularies, metadata and data standards for maximized Integration, see Strategic Pathway 6: Standards.

4.7 Deliverables

The list of deliverables below are the outcomes typically created as a result of completing the actions in this strategic pathway. They are key success indicators in realizing an Integrated Geospatial Information Framework. Examples include:

- User requirement
- Data Framework
- Data Inventory
- Data Profiles, Theme data specifications and Metadata
- Data Gap Analysis
- Data Theme Road Map
- National Data Capture and Acquisition Program
- Data Custodianship Policy and Guidelines
- Data Governance and Management
- Data Release, Storage and Retrieval Processes
- Maintained Geodetic Infrastructure
- Geospatial and Statistical Integration
- Geocoding and Aggregation
- Data Supply Chains and Interoperability

4.8 Outcomes

The following outcomes result from the execution of well-defined data supply chains and custodianship guidelines throughout the geospatial data life cycle:

- An increased range and scope of authoritative, integrated geospatial information available as insights for decision-making and policy-setting to address economic, social and environmental challenges.
- A critical mass of centrally coordinated data discoverable to support national development and innovation leading to economic growth and improved quality of life for citizens.

- Cost reduction through productivity improvements achieved via well-defined supply chains that eliminate duplication and make standardized data accessible to end users for integration and reuse.
- Ability to monitor and measure progress towards achieving broad socio-economic benefits, including the sustainable development goals, through access to quality geospatial information.

4.9 Resources

As part of the work programme of UN-GGIM, there are a number of related initiatives and activities including by the Subcommittee, Expert and Working Groups of the Committee of Experts. These initiatives and activities are multi-stakeholder when arriving at outcomes and outputs. This inclusive and participatory nature of work has allowed the preparation of a number of resource documents/publications that are helpful and useful when addressing the complexities in data that impact geospatial information management. This includes specifically the work and contributions of the UN-GGIM Sub-Committee on Geodesy, the Working Group on the Global Fundamental Geospatial Data Themes and the Expert Group on the Integration of Statistical and Geospatial Information. Their work and adopted frameworks have provided a series of deliverables that will support countries in developing their data methodologies for geospatial information management and have been used in this strategic pathway. These include:

- The Global Geodetic Reference Frame (GGRF) Roadmap[9] and Implementation Plan[10] developed by the Sub-Committee on Geodesy;
- The Global Fundamental Geospatial Data Themes[11]; and
- The Global Statistical Geospatial Framework GSGF[12]).

4.10 References

Arnold (2016) Improving Spatial Data Supply Chains: Learnings from the Manufacturing Industry, GeoProcessing 2016: The Eighth International Conference on Advanced Geographic Information Systems, Applications & Services, Venice, Italy. IARIA.

ANDS (n.d.) Fair Data Principles [Online] Available at https://www.ands.org.au/working-with-data/the-fair-data-principles, accessed January 2020.

Cooperative Research Centre for Spatial Information, CRCSI (2017) Towards a Spatial Knowledge Infrastructure [Online] Available at https://www.crcsi.com.au/assets/Program-3/CRCSI-Towards-Spatial-Knowledge-Whitepaper-web-May2017.pdf

[9] GGRF Road Map available at http://ggim.un.org/meetings/GGIM-committee/documents/GGIM6/E-C20-2016-4%20Global%20Geodetic%20Reference%20Frame%20Report.pdf

[10] GGRF Implementation Plan available at http://ggim.un.org/meetings/GGIM-committee/8th-Session/documents/Road-Map-Implementation-Plan.pdf

[11] The Global Fundamental Geospatial Data Themes, available at: *http://ggim.un.org/meetings/GGIM-committee/9th-Session/documents/Fundamental_Data_Publication.pdf*

[12] GSGF available at http://ggim.un.org/meetings/GGIM-committee/9th-Session/documents/The_GSGF.pdf

Data Interoperability Standards Consortium, DISC (2017), [Online] available at http://datainteroperability.org/, accessed January 2020.

Geospatial World (2010) Metadata for national Spatial Data Infrastructure [Online] Available at https://www.geospatialworld.net/article/metadata-for-national-spatial-data-infrastructure/, accessed January 2020.

Nebert (2004). SDI Cookbook, Version 2.0, [Online] Available at http://gsdiassociation.org/images/publications/cookbooks/SDI_Cookbook_GSDI_2004_ver2.pdf, Accessed January 2020.

UN-GGIM (2015) Future trends in geospatial information management: the five-to-ten-year vision, Second Edition, [online] available at: https://ggim.un.org/documents/UN-GGIM-Future-trends_Second%20edition.pdf

Yu, F., West, G., Arnold, L., McMeekin, D., & Moncrieff, S. (2016) "Automatic geospatial data conflation using semantic web technologies". In ACM International Conference Proceeding Series Vol. 01-05-February-2016.

Van Oosterom, P. and Ziatanova, S. (2008) Creating Spatial Information Infrastructures: Towards the Spatial Semantic Web, CRC Press, page 156, Taylor and Francis Group

Wilkinson, M. D., Dumontier, M., Aalbersberg, I.J., Appleton, G., et al. (2016) The FAIR Guiding Principles for scientific data management and stewardship. Scientific Data. 3: 160018. doi:10.1038/sdata.2016.18. OCLC 961158301. PMC 4792175. PMID 26978244