The Global Statistical Geospatial Framework

Second Edition

2025

Department of Economic and Social Affairs

Statistics Division

Secretariat of the United Nations Committee of Experts on Global Geospatial Information Management

The Global Statistical Geospatial Framework Second edition

United Nations
New York, 2025

Department of Economic and Social Affairs

The Department of Economic and Social Affairs of the United Nations Secretariat is a vital interface between global policies in the economic, social, and environmental spheres and national action. The Department works in three main interlinked areas: (i) it compiles, generates and analyzes a wide range of economic, social and environmental data and information on which Member States of the United Nations draw to review common problems and to take stock of policy options; (ii) it facilitates the negotiations of Member States in many intergovernmental bodies on joint courses of action to address ongoing or emerging global challenges; and (iii) it advises interested Governments on the ways and means of translating policy frameworks developed in United Nations conferences and summits into programs at the country level and, through technical assistance, helps build national capacities.

Statistics Division

The United Nations Statistics Division (UNSD) is committed to the advancement of the global statistical and geospatial systems. It compiles and disseminates global statistical information, develops standards and norms for statistical and geospatial activities, and supports countries' efforts to strengthen their national statistical and geospatial information systems. UNSD facilitates the coordination of international statistical and geospatial activities and supports the functioning of the United Nations Statistical Commission, the United Nations Committee of Experts on Global Geospatial Information Management, and the United Nations Group of Experts on Geographic Names as the apex entity of the global statistical and geospatial systems.

United Nations Committee of Experts on Global Geospatial Information Management

The United Nations Committee of Experts on Global Geospatial Information Management (UNGGIM) is the apex intergovernmental body to discuss, enhance and coordinate global geospatial information management activities by involving Member States at the highest level, to work with Governments to make joint decisions and set directions on the use of geospatial information within national and global policy frameworks, and to develop effective strategies to build geospatial capacity in developing countries.

United Nations Statistical Commission

The United Nations Statistical Commission, established in 1947, is the highest body of the global statistical system. It brings together the Chief Statisticians from member states from around the world. It is the highest decision-making body for international statistical activities especially the setting of statistical standards, the development of concepts and methods and their implementation at the national and international level. The Statistical Commission oversees the work of the United Nations Statistics Division (UNSD) and is a Functional Commission of the UN Economic and Social Council.

Notes

The designations used and the presentation of material in this publication do not imply the expression of any opinion whatsoever on the part of the Secretariat of the United Nations concerning the legal status of any country, territory, city or area, or of its authorities, or concerning the delimitation of its frontiers or boundaries. The term “country” as used in this publication also refers, as appropriate, to territories or areas. The designations “developed regions” and “developing regions” are intended for statistical convenience and do not necessarily express a judgment about the stage reached by a particular country or area in the development process.

Foreword

In an era where data-driven decision-making is more critical than ever, the integration of statistical and geospatial information plays a fundamental role in shaping policies that impact economies, societies, and the environment.

The United Nations Global Statistical Geospatial Framework (GSGF) provides a structured approach to linking statistical and geospatial data, enabling National Statistical Offices (NSOs) to effectively engage with National Geospatial Information Agencies (NGIAs) allowing governments worldwide to produce more accurate, consistent, locally relevant, and actionable national statistics. This integration is vital for evidence-based policymaking, supporting the monitoring of sustainable development, disaster response, urban planning, and countless other areas where location intelligence enhances decision-making.

Since the first edition of the GSGF, the need for a more accessible, streamlined, and implementable version has become clear. While the principles remain robust, experience has shown that simplification is key to broader adoption, particularly at the country level.

This second edition of the GSGF responds to that need. It builds on the successes of the first edition but has been refined to ensure that implementation is easier, more intuitive, and aligned with real-world national statistical processes. By making the framework more user-friendly, we aim to empower countries to harness the power of integrated statistical and geospatial data, unlocking new insights that drive sustainable development and improved governance worldwide.

We thank member states from the expert group for their support in producing this second edition, and the continued efforts in creating implementation guides and supporting resources to complement this document.

We encourage NSOs, policymakers, and geospatial communities to embrace this framework and leverage its principles to build stronger, more resilient, and more data-driven societies.

Mr Claudio Stenner
Brazilian Institute of Geography and Statistics (IBGE)
Brazil

Ms Olivia Powell
Office for National Statistics (ONS)
United Kingdom

Co-Chairs
United Nations Expert group on the Integration of Statistical and Geospatial Information

Contents

Executive Summary

In today's data-centric world powered by ever faster evolving technology, the integration of statistical and geospatial information is crucial for informed decision-making across various sectors.

Antonio Guterres
Secretary General
of the United Nations (2018)

> *"Reliable, timely, accessible and disaggregated geospatial information must be brought to bear to measure progress, inform decision-making and ensure effective and inclusive national and sub-national programs that will chart the path toward the 'Geospatial Way to a Better World'"*

The **Global Statistical Geospatial Framework** (GSGF) serves as a standardized and harmonized approach to the integration of data, enhancing the quality and accessibility of geospatially enabled statistics.

Through the application of its five Principles, the GSGF facilitates the consistent production and integration of geospatial and statistical information, enabling enhanced decision-making through streamlined processes, reducing redundancy and improving the efficiency of data collection and analysis. By combining location-based data with statistical analysis, policymakers can address complex issues such as urban planning, public health, and environmental management more effectively. It also provides a unified framework which allows for standardized data integration, promoting intra- and inter-national comparability.

The GSGF is built upon five guiding Principles:

1. **Use of Fundamental Geospatial Infrastructure and Geocoding:** Establishes a common approach to accurately assign location references to statistical units, ensuring precise data integration using location.
2. **Geocoded Unit Record Data in a Data Management Environment:** Advocates for the management of detailed geocoded data within secure environments, facilitating robust data analysis while maintaining confidentiality.

3. **Common Geographies for Dissemination of Statistics:** Promotes the use of standardized geographical units to ensure consistency and comparability in data dissemination and analysis.
4. **Statistical and Geospatial Interoperability:** Establishes interoperability of data, systems, and processes through the implementation of internationally adopted standards and best practices.
5. **Accessible and Usable Geospatially Enabled Statistics:** Ensures that integrated data is both accessible and usable for a wide range of stakeholders, supporting transparent and informed decision-making.

Governance and institutional capacity, policy and legal, human resources and capability, data and interoperability, and technology and infrastructure are all key themes which are supporting the GSGF Principles. Maturity in these areas will help accelerate the implementation of the GSGF Principles.

There are clear and tangible benefits of implementing the GSGF: it will improve quality, speed, and efficiency of national statistical operations like censuses and surveys by cross-referencing new and extensive sources of geospatial/location-based information, allowing for innovative and more localized insights. This in turn allows National Geospatial Information Agencies to access a wider variety of training data to integrate into the creation of high-quality inputs for use in small area estimates and demographic modeling efforts. Consequently, this enables better decision-making by offering analyzes and indicators that address increasingly complex national and international policy questions.

Conversely, failing to utilize the GSGF can lead to significant operational risks: unnecessarily costly data collections, inefficient statistical processes, poor reproducibility compounded by the inability to ingest new, valuable datasets, and lower quality data products due to outdated, stagnant methods of data production. These deficiencies maintain outdated data production methods, resulting in diminished understanding of national and local challenges, poor resource allocation, and reduced capacity for response during critical events such as pandemics or natural disasters.

Implementation of the GSGF facilitates the development of common, standardized tools supporting data integration and data sharing at international, national, and sub-national levels. By providing a robust data infrastructure for monitoring progress and making informed decisions, including and especially during crises, the framework strengthens efforts to achieve the Sustainable Development Goals (SDGs) at the national, regional, and global scale.

By adopting the GSGF, nations can harness the power of integrated statistical and geospatial data, leading to more informed policies and a deeper understanding of complex societal challenges to resolve issues in the location where they are best addressed.

List of Key Abbreviations

API	Application Programmable Interface
DGGS	Discrete Global Grid System
EG-ISGI	The (United Nations) Expert Group on the Integration of Statistical and Geospatial Information
FAIR	Findable, Accessible, Interoperable, Reusable Principles
GIS	Geographic Information Systems
GSBPM	Generic Statistical Business Process Model
GSGF	Global Statistical Geospatial Framework
GSIM	Generic Statistical Information Model
HLG-MOS	(United Nations) High–Level Group for the Modernization of Official Statistics
IGIF	Integrated Geospatial Information Framework
IHO	International Hydrographic Organization
ISO	International Standards Organization
MoU	Memorandum of Understanding
NGIA	National Geospatial Information Agency[1]
NSDI	National Spatial Data Infrastructure
NSO	National Statistical Organization
NSS	National Statistical System
OGC	Open Geospatial Consortium
SDGs	Sustainable Development Goals
SDMX	Statistical Data and Metadata eXchange
UNECE	United Nations Economic Commission for Europe
UNGGIM	The (United Nations) Committee of Experts on Global Geospatial Information Management
UNSC	United Nations Statistical Commission
UNSD	United Nations Statistics Division
WFS	Web Feature Services
WMS	Web Mapping Services
XML	Extensible Markup Language

[1] NGIA is used as an encompassing term to cover National Mapping, National Cartographic, National Geospatial Information Agencies and Authorities.

Introduction

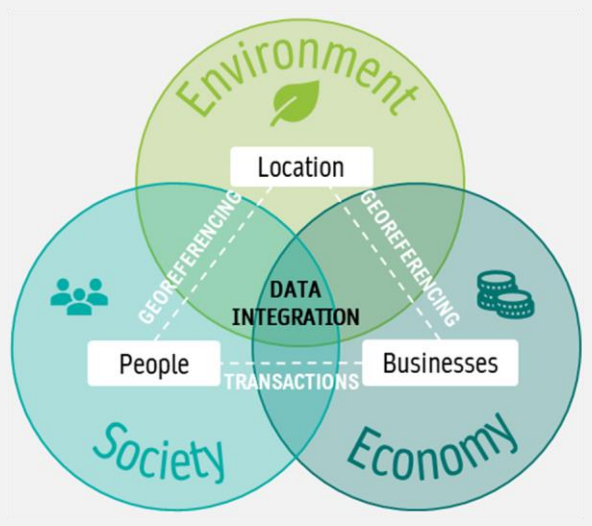
We live in a data rich world, with fast evolving technology fueled by a growing appetite for data insight. With the increase of the availability of highly accurate data both in terms of timescale (real-time data) and geographical location (granularity), managing and integrating data can be difficult.

National and local governments rely heavily on administrative, geospatial, and other non-statistical data to produce statistical outputs and coupled with the growing need for localized information, the importance of integrating statistical data with geospatial data has never been so critical.

What is Geospatial Information?

Geospatial information is data which describes objects, events or other features and their location on earth. This can include addresses and geographical areas, foundational base mapping (e.g., from NGIAs), environmental data and increasingly, earth observations and sensor data such as mobile phone data.

The 2030 Agenda for Sustainable Development[2] creates an imperative to understand, sustain, and develop the three pillars for development: **Economy, Society, and Environment**[3]. Positive outcomes in these three areas are maximized by enabling data-driven, evidence-based decision making within both the public and private sectors. The 17 Sustainable Development Goals[4] (SDGs), 169 targets and global indicator framework of the 2030 Agenda all depend on sensitive, precise, and comparable data and measurements.

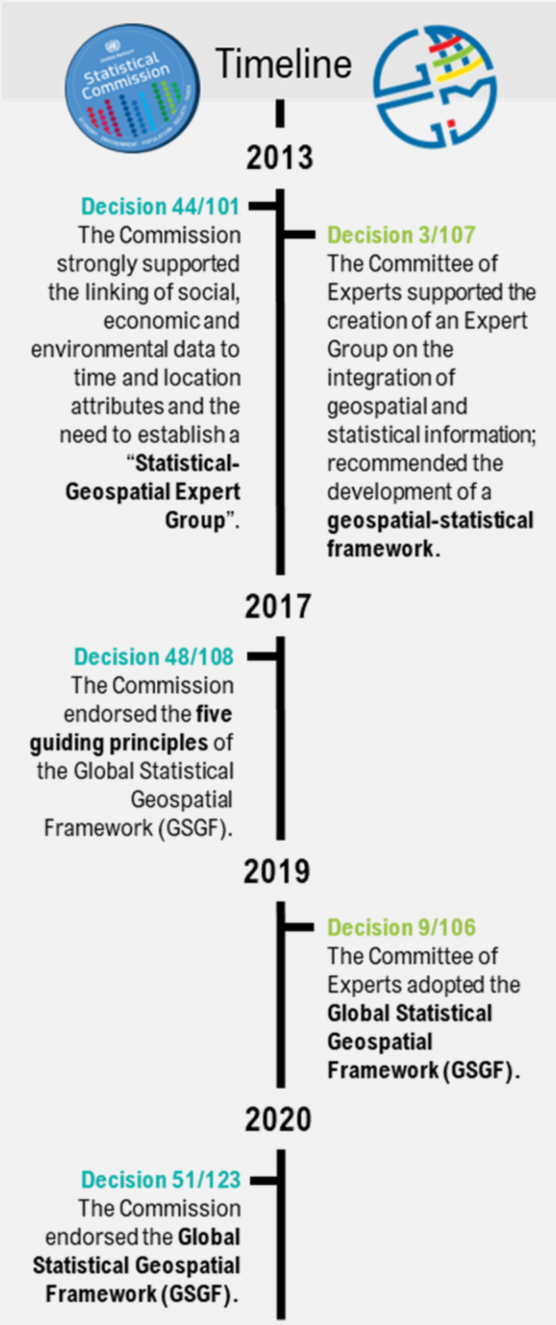


Source: UNGGIM, UNSD (2025)

National statistical measures center around those same three pillars. They require both **statistical** (i.e., socio-economic data) and **geospatial** data (i.e., location or position information, earth observation, environmental, and mobility data). **Geography** is the common language or bridge between disparate datasets about people, businesses, and the environment we live in. **Location** is a critical attribute to statistical data and the most efficient way to combine existing datasets to the vast amount of emerging geospatial datasets available.

[2] UN Sustainable development Goals web page: https://www.un.org/sustainabledevelopment/development-agenda/
[3] https://www.un.org/ecosoc/en/sustainable-development
[4] https://www.un.org/sustainabledevelopment/sustainable-development-goals/

The Mandate



Source: UNGGIM, UNSD (2025)

About the GSGF

The **Global Statistical Geospatial Framework (GSGF)** is a high-level framework adopted both by the United Nations' Statistical Commission (UNSC) and the Committee of Experts for Global Geospatial Information Management (UNGGIM).

It provides guidance to NSOs across the globe on the use of geospatial data for the purpose of enhancing statistical production. It provides the underlying mechanism to achieve data integration, delivering geospatially enabled statistics which provide a modern and complete picture of our world.

It bridges the gap between the two well-known United Nations' (UN) frameworks for statistical processes and for geospatial information:

- the Generic Statistical Business Process Model (GSBPM)[5] and
- the UN Integrated Geospatial Information Framework (UN-IGIF)[6]

Developed by the Expert Group on the Integration of Statistical and Geospatial Information (EG-ISGI), it consists of five Principles to achieve successful data integration and deliver geospatially enabled statistics.

[5] United Nations, Department of Economic and Social Affairs, Statistics Division (2019), GSPBM v5.1.
[6] United Nations (2023), United Nations Integrated Geospatial Information Framework Part1.

How to use this paper

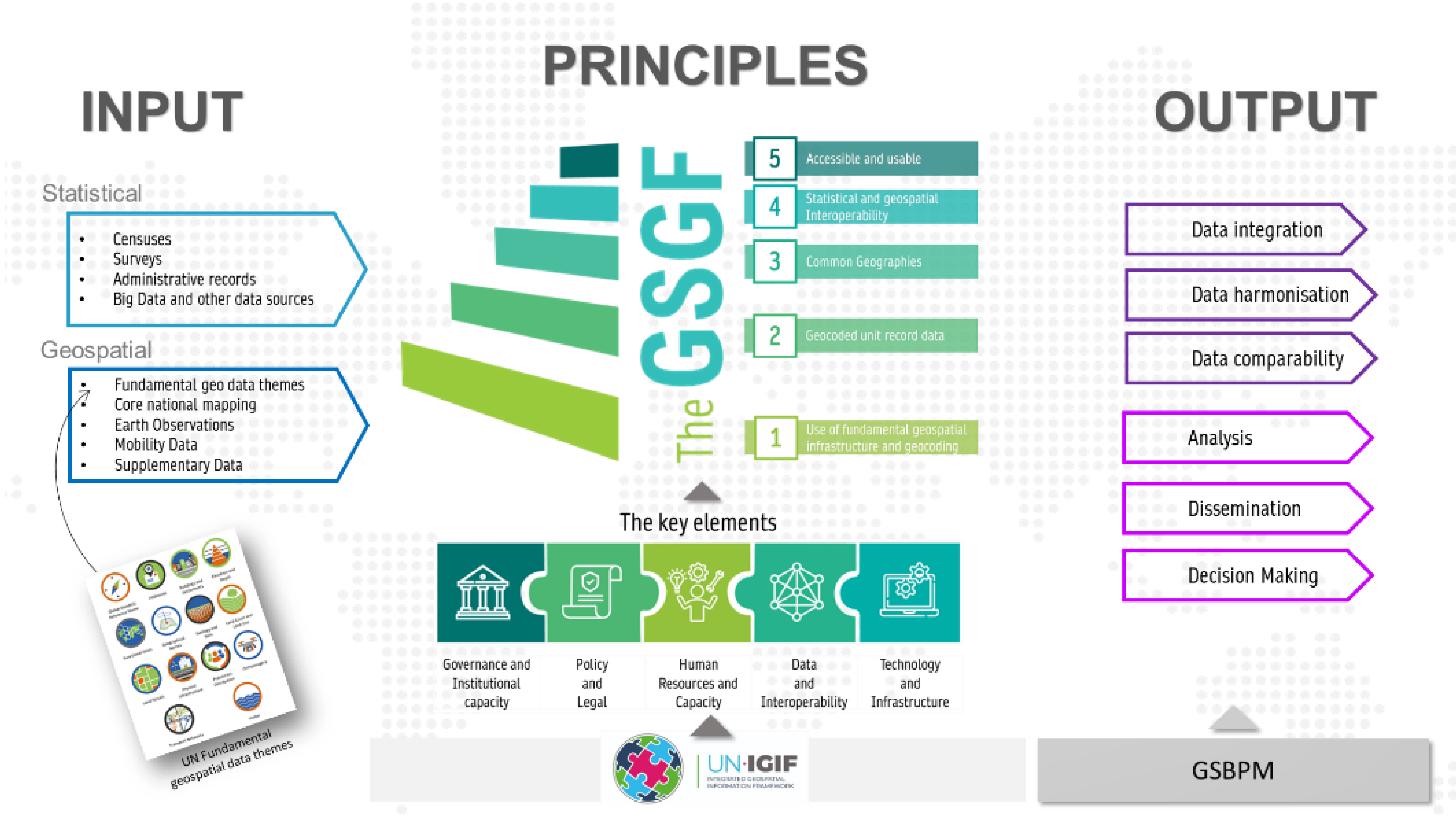
This document serves as a reference document and introduction to the GSGF, focusing on high-level requirements and principles needed to foster good geo-statistical data integration.

- Part One presents the framework, its context and links to other frameworks.
- Part Two provides an in-depth overview of each of the five Principles and their requirements, including the Key Elements necessary to support their development and implementation.

The EG-ISGI is the custodian of the GSGF and will release materials to further support countries in their implementation of the five Principles through guidance, good practices, standards development as well as case studies and lessons learned from national and regional implementations.

Implementation guides sitting alongside the GSGF already exist. Further resources to strengthen capability building, implementation and operationalization of the GSGF will be added to the UN Wiki, aiming to provide “living documents” to keep up-to-date and reflect new technological advancements and data practices, including detailed techniques, case studies, and other relevant guidance.

PART ONE: The GSGF framework and its components



Source: UNGGIM, UNSD (2025)

Part 1: The GSGF framework and its components

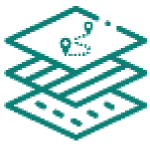
The GSGF is a four-stage framework covering data inputs, Key Elements, the five Principles, and outputs. Through these components the GSGF serves as a bridge between organizations, standards, methods, workflows, tools in both the statistical and geospatial domains.

Stage 1: Data Inputs

As inputs, the GSGF includes data coming from both the geospatial and statistical domains.

Statistical data input refers to official statistics either produced by a governmental body (e.g., censuses and surveys), or produced through other methods but certified as official by a governmental body. These datasets are increasingly drawn from administrative records and other official sources. These datasets come from NSOs and administrative data custodians within the broader National Statistical System (NSS), e.g., across government and public sector bodies. These data sources are often augmented by complementary datasets, such as those derived from crowdsourcing, new technology or other sources (e.g. Big Data).

Usually, this information comes in the form of tabular datasets and can also include some kind of location or "place" information as an attribute e.g., residential address, postal codes, place name, etc. which can be used to perform data integration to other datasets.



Geospatial Data input refers to the foundational geospatial data which provides the human and physical geographical context of a country. It also includes new sources of data such as sensor data e.g., earth observations and mobility data, which are increasingly used to detect built-up extent, population movement, etc. Geospatial data inputs also include any geospatial data used in analysis and linked to other variables, such as geographical codes or coordinates. The biggest challenges we face today such as environmental pollution, land use, and climate change, all are inherently tied to a specific location and have heterogenous impacts across our landscape. e.g., air pollution, land use, climate data. In this context, geospatial data usually serves to geospatially enable authoritative statistical data and increasingly contribute to statistical production.

These traditional geospatial datasets are primarily provided in both vector and raster formats, depending on the source and method of dissemination, and are increasingly provided as digital network-based data services, allowing machine to machine access.

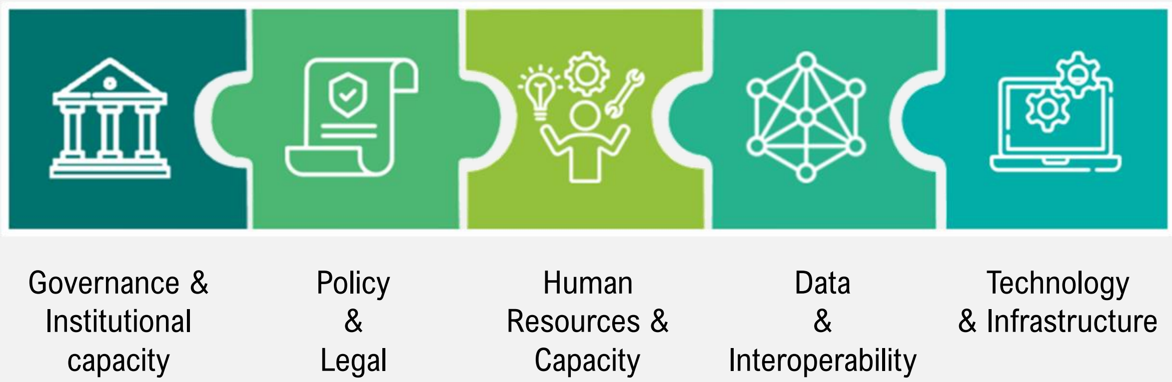
As geospatial data production becomes more accessible and data sources proliferate, NSOs must develop strategies for the integration of data from the broader geospatial community and non-traditional sources, including private industry, crowdsourced datasets, and non-governmental organizations.

Stage 2: The Key Elements

The GSGF Key Elements refer to the essential institutional and national infrastructure that enables implementation of the GSGF and its Principles. They align to the nine pathways[7] of the **UN-IGIF[8] framework** but focus on those relevant to NSOs for the purpose of statistical and geospatial data integration.

Following the publication of the UN-IGIF, the GSGF Key Elements from the first edition were reviewed and aligned to the nine strategic pathways of the UN-IGIF. Previously known as 'the four Key Elements,' "Human Resources and Capacity" was added as a "fifth" element in recognition of the importance of a skilled workforce in the creation of geospatially enabled statistics.

The GSGF Key Elements



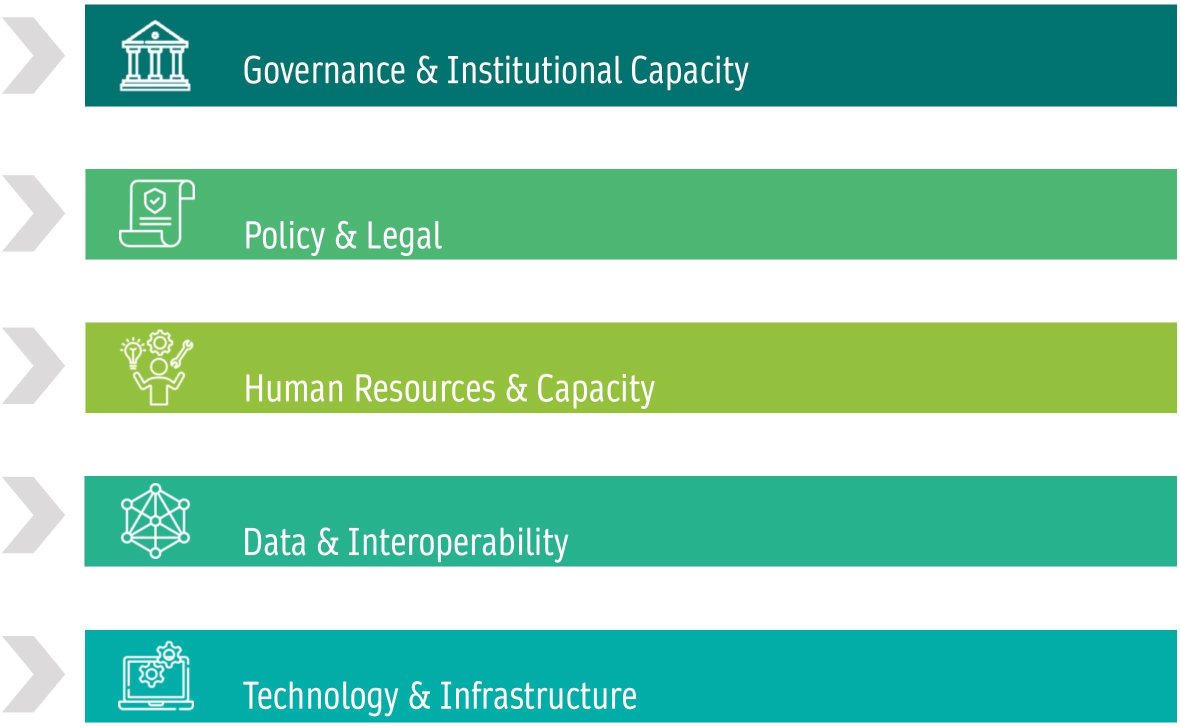
Source: UNGGIM, UNSD (2025)

These are the five topical areas, components of an ecosystem, which are the support infrastructure and critical to the development of a fully integrated geospatial and statistical data ecosystem. Implementation of the GSGF is not an endpoint, it is a process, and it is recommended to adopt an iterative process to promote continuous improvement over time.

The GSGF Five Principles and Key Elements Matrix in Appendix 4 summarizes the extent to which each Principle depends on the various Key Elements.

Explicit requirements are detailed in Part 2: The Five GSGF Principles under each Principle section – clearly referencing the associated Key Element specific requirements using the following structure.

[7] See more details about the UN-IGI and its nine pathways in Appendix 2.
[8] UN-IGIF web page: https://ggim.un.org/UN-IGIF/



To support the implementation of the framework, the EG-ISGI has developed a GSGF maturity self-assessment tool[9], adapted from the UN-IGIF baseline assessment methodology and Diagnostic Template[10]. By presenting targeted questions aligned with the GSGF Key Elements, this tool enables NSOs to conduct comprehensive assessments of their strengths and weaknesses across each Key Element. The resulting analysis empowers nations to develop prioritized GSGF action plans focused on critical improvement areas.

[9] UN Statistics Wiki: UN EG-ISGI GSGF Self-Assessment Tool (Excel file)

[10] The World Bank UN-IGIF Baseline Assessment Methodology (PDF) and Diagnostic Template (XLSX) are available here: https://www.wbgkggtf.org/node/3547

Stage 3: The Five Principles of the GSGF

The Five Principles of the GSGF outline the broad processes by which a range of geospatial and statistical infrastructures and processes are applied following the data input stage, through to integration. Firstly, statistical data are geospatially enabled or linked using the more precise and accurate location data available to spatially tie the statistical data to a specific location. Then, geospatial tools and methods, such as common geographies and common standards of good practice, are used to ensure the data are interoperable, accessible, discoverable, and usable.

The Five GSGF Principles

Source: UNGGIM, UNSD (2025)

These Principles guide countries in establishing and strengthening their processes for geospatially enabling statistical and administrative data and allow for the identification of capacity gaps within their respective national statistical and geospatial infrastructure.

More information about each Principle can be found in Part 2: The Five GSGF Principles.

Stage 4: GSGF outputs

The GSGF aims to provide a structured approach to linking statistical and geospatial data to avoid disparate silos of geospatially enabled datasets with consistent results, to create analysis ready datasets.

Effective implementation of the GSGF Principles, grounded by the Key Elements, can provide many positive outcomes, including the incorporation of spatial linkages into the data production workflow, resulting in fully integrated primary data streams. This creates a higher degree of structural harmonization and standardization facilitating analysis over time, space, and geographic units. This enhanced data environment at the national or enterprise level provides:

- **Data Integration:** Heterogenous datasets can be combined using location as the primary key or common denominator even at localized/granular levels. This allows for better linked data at the enterprise level, and a greater ability to produce novel analyzes bringing disparate datasets together.
- **Data Harmonization:** Datasets are consistent in definitions, formats, and classifications, facilitating and streamlining data processing and cleansing, for better analysis results.
- **Data Comparability:** Datasets can be compared at various geographical levels and through time, facilitating longitudinal studies and regional and international analysis.

These enhanced outputs based on datasets with geospatial identifiers built into the initial production workflow have an inherently greater capacity for location-based data integration compared to methods that join statistical and geospatial data as a post-process at the end of their respective production pipelines, or by data users, if at all. This integration enhances the ability to create reproducible results across comparable units of observation and at various temporal and geographic scales. These outputs support more efficient and powerful **analysis**, **dissemination** of information and applications that support informed **decision-making** processes.

Building geospatially enabled statistical data during the initial production phase also facilitates more complex statistical and geospatial **analyzes** – spatial statistics can disaggregate record data into locally meaningful resolutions through areal interpolation techniques such as dasymetric modeling or provide more flexibility with greater aggregation options.

It supports many national data processes and policy applications such as (but not only):

Social	Economic	Environmental
• population census • household census and surveys • population registers • population migration	• business registers • agriculture census • employment or labor markets • price surveys • tourism	• natural capital • climate change

PART TWO: The Five GSGF Principles

Principle 1: Use of fundamental geospatial infrastructure and geocoding

1

Principle 1 focuses on the creation of geospatial infrastructure which supports the accurate assignment of location information to data records at the national level, bringing together relevant stakeholders and partners to ensure the development of a high-quality and consistent national location referencing framework (such as a physical address, property or building identifiers or other location descriptions).

? Definition: What does this Principle cover?

Principle 1 recommends the development of a **National Spatial Data Infrastructure** (NSDI)[11] supporting the provision of core geographic information at the national level. This is in line with the UN-IGIF under "Technology". This is a national endeavor where one or more organizations are responsible for the infrastructure.

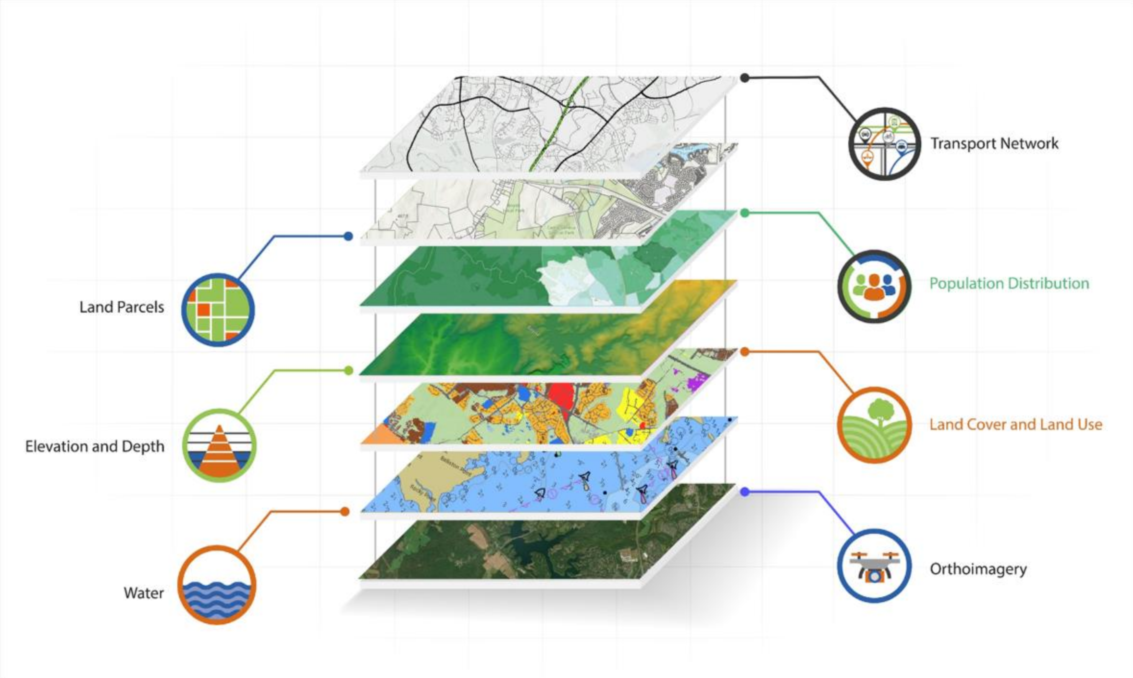
The UNGGIM Global Fundamental **Geospatial Data Themes**[12] provides a list of essential datasets, relating to physical or human geographical features, grounded in a robust coordinate reference system describing the world we live in. These datasets aim to provide the basis of spatial location data references that will be used in other Principles.

These datasets might come in various geospatial "layers" – see concept diagram overleaf. Geospatial layers are a way to organize and represent different types of geographic information in a Geographic Information System (GIS). They are essentially digital overlays that can be combined to create a comprehensive map or analyze spatial relationships.

Having the national infrastructure in place enables countries, implementers, and data users to accurately **geo-reference** statistical data at the microdata/unit level (e.g., a person, household, business, building or parcel/unit of land) by associating each statistical record with a corresponding location and time of collection.

[11] For a definition of NSDI, please refer to Appendix 6: Definitions of Terms
[12] UNGGIM (2019), The Global Fundamental Geospatial Data Themes publication

The concept of geospatial layers



Source: UNGGIM, UNSD (2025)

Objectives and Benefits: Why do we need this Principle?

To enable the production of high-quality, persistent, standardized, and trusted national statistics, it is necessary to have a common and consistent approach to establishing the accurate location of each unit in a statistical dataset. If the intent is to count a person (or other statistical unit) once and only once, and in the right location, it's necessary to be able to assign a unique identifier to that statistical unit and record the time and place it was counted to avoid duplication.

A national geospatial infrastructure provides the geospatial data themes and reference frames needed for statistical operations and acts as the foundation for the delivery of robust statistical business processes through greater integration of statistical and geospatial information. It charts the course for geographically referencing the data to be collected, as accurately and consistently as possible, starting from the earliest design stages. The Geospatial Generic Statistical Business Process Model[13] (GeoGSBPM) incorporates "location" or "place" as a theme

[13] GeoGSBPM - Generic Statistical Business Process Model - UNECE Statswiki

throughout the eight stages of the recommended statistical business practices, to answer the question of "where" a statistical data point was collected.

The main goal of Principle 1 is to ensure a resilient, fundamental geospatial infrastructure for statistical and geospatial data exists, which leads to achieving the following objectives:

- **Infrastructure:** Collaborate with national geospatial stakeholders to establish a geospatial solution supporting national needs.
- **National Location Referencing Framework**: Address, property, building, and location information are accurate and consistent, meeting country-level agreed standards and good practices. It would add data coherence between national statistics and geospatial data.
- **Data Quality:** Geocoding results are as accurate and consistent as possible using common approaches or systems. Any geocoding issues are consistently managed through the application of standardized approaches.

Implementing Principle 1 will bring the following benefits[14]:

Strategic Benefits	Benefits for NSOs
• Greater stakeholder collaboration at national level, leveraging national fundamental data provision and the adoption of standardized data themes. • Reduction in the risk of duplication in the production of core fundamental geospatial datasets at the national level, thus driving efficiencies. • Diverse inputs into national geospatial information requirements, which are responsive to policy making across environmental, societal, and economic domains. • Greater data interoperability and integration between geospatial stakeholders, following agreed national frameworks.	• Maintained national lifecycle information of geospatial features for consistent linking of statistics and geographic location information. • Enhanced reproducibility and quality of statistical results. • Enabled provision of new geospatial information sources for measuring and monitoring targets and indicators, such as the SDGs. • Reduction in the resources needed to conduct censuses and surveys by providing the location information needed for address frames and enumeration areas.

[14] Answers compiled from a workshop at the Expert Group for the Integration of Statistical and geospatial Information in Nairobi in September 2024 and combined with existing published benefits.

Key Stakeholders: Who needs to be involved?

The provision of fundamental geospatial datasets is a responsibility shared across a wide range of institutions from either the geospatial or the statistical domains, although varying from country to country. Nevertheless, NSOs will need to ensure they have access and are using the relevant, authoritative data themes and must pursue collaborative engagements with the responsible institutions.

Common stakeholders will include: NGIAs; relevant government institutions / ministries; NGOs and civil societies; geodetic communities; academia and scientists; private enterprises; citizens (e.g., crowdsourcing); and data suppliers.

Requirements: What does this Principle need to be implemented?

Principle 1 mostly relates to the data requirements. It prescribes the availability and use of the following six (as a minimum) of the fourteen UNGGIM Global fundamental geospatial **datasets**, at national level, with their applied date and time stamp[15]:

- Geodetic Reference Frames
- Grid or coordinate system
- Address register, property or building identifiers
- Land parcels or cadastre
- Place / Geographical names
- Set of functional areas

Governance & Institutional Capacity

Each country should have an institution responsible for the national geodetic[16] service, usually the national mapping or geospatial information agency. Geodesy is the center to support the localization of key physical features which form the base of the fundamental datasets.

[15] Inclusion of a time and date stamp, bringing a temporal component to the geospatial and statistical unit record data. This is particularly important as geographical boundaries and our physical environment changes over time, and this needs to be reflected in the positional data.
[16] Refer to Appendix 0:

Definitions of Terms for a definition of 'geodesy' or 'geodetic service'.

These datasets are the responsibility of national institutions or stakeholders. This will vary depending on the national landscape and geospatial ecosystem. NSOs will have a strong dependency on these geospatial datasets, however may lack the responsibility and authority over them. Therefore, it is necessary to identify and foster **strong collaborative relationships** between NSOs, key national stakeholders, and providers of the core geospatial datasets.

Policy & Legal

Establishing **strong communication** and **institutional collaboration mechanisms** between NSOs and NGIAs is essential. This can be facilitated by, for example, country-level laws and policies, Memorandum of Understandings (MoUs)[17], data sharing agreements, and other communities of practice. It is recommended to scope the data sharing landscape in place at country level to support discussions and establish the needs for new arrangements.

Human Resources & Capacity

It is important to build **data literacy skills** which include geospatial data, so staff at all levels of government organizations, including NSOs, understand the specificities of geospatial data and what the core datasets are.

To be able to collect, store and manipulate the core datasets, proficiency with geospatial technology and geospatial data management is key. Often, NSOs will have a **'GIS' or 'Geography' team**, providing geospatial data and support to the various business areas.

For the technology infrastructure required to host and make available the UN fundamental geospatial data themes, specialized geospatial solution architectures knowledge is required – see Technology and Infrastructure section below.

Data & Interoperability

The geospatial datasets stated above, should be from comprehensive, **authoritative** data sources, adhering to open data standards whenever applicable. This ensures the quality, accuracy, currency, and consistency of data from data custodians within the NSS and NSDI thereby helping to meet country-level priorities, align with accepted international agendas, and adhere to internationally recognized standards and good practices.

[17] For an example of an institutional arrangement, refer to the United Nations' MoU template for joint work between the NSO and the NGIAs for the integration of statistical and geospatial information: https://ggim.un.org/meetings/GGIM-committee/14th-Session/documents/Background_document_Institutional_agreement_for_GSGF.pdf

Technology & Infrastructure

NSDIs enable the effective production, storage and dissemination of the UN's 14 global fundamental geospatial data themes (or an expanded set of national geospatial data assets). A NSDI "is a means to assemble geographic data nationwide to serve a variety of users"; The UN-IGIF recommends such a framework as it facilitates the efficient integration, management, and maintenance of spatial data for a country. It is meant to be comprehensive, standards-based technology infrastructure built on four foundational pillars: robust hardware systems, interoperable software platforms, reliable networking capabilities, and standardized data management protocols.

NSDIs vary globally with regards to their national implementations. Some countries may have one agency in charge of the NSDI collating all fundamental geospatial data themes in one infrastructure (spatial database); others may operate under an ecosystem of agencies sharing responsibilities in the custodianship of the geospatial data themes. Whether the system is on-premises, cloud-based, or hybrid, the infrastructure should support distributed data management where agencies maintain authoritative control while participating in federated sharing. NSDI will require expert geospatial technology solution architecture knowledge to be built and sustained.

NSDIs must be designed for future evolution, incorporating real-time sensor integration, machine learning for automated quality assurance, mobile-first interfaces for field data collection, and seamless integration with existing frameworks. Success depends on implementing governance frameworks that coordinate technical standards adoption, support capacity building programs, and establish sustainable funding models for long-term maintenance and technological advancement.

NSOs will need to share their requirements to the NSDI custodian agency in regard to accessing, using or depositing geospatial datasets. This might include hardware, software architecture which implement a service-oriented approach following international standards, enabling seamless data sharing through standardized web services.

NSOs will require its own geospatial technology infrastructure to be "compatible" with the NSDI and should be consulted in its design to ensure the infrastructure supports statistical production lifecycles, including decennial censuses, household surveys, and the integration of administrative and geospatial data sources. For NSOs, essential geospatial infrastructure components include:

- Spatial database management systems supporting 2D/3D geometries and statistical data integration.
- Support for robust geocoding and georeferencing, development of spatial sampling frames, and linkage of administrative and survey data to geospatial datasets.

- Automated workflows for maintaining data currency and quality across all themes.
- Comprehensive metadata management following international standards for resource discovery.
- API frameworks enabling integration between agency systems and international data sharing networks.

Links and Resources

Below are links to further reading or resources relating to Principle 1:

- The Global Fundamental Geospatial Data Themes
- ISO/TC 211 Addressing pages
- ISO/TC211 Geodetic Registry
- United Nations Global Geodetic Centre of Excellence
- United Nations Expert Group on Land Administration and Management
- United Nations MoU template for joint work between the NSO and the NGIA for the integration of statistical and geospatial information

National examples and NSDIs implementations:

- Australian Spatial Data Infrastructure (ASDI) - https://www.icsm.gov.au/australian-spatial-data-infrastructure-asdi
- European GISCO - https://ec.europa.eu/eurostat/web/gisco
- National Spatial Data Infrastructure of India
- NSDI Act of Japan - https://www.gsi.go.jp/kokusaikoryu/kokusaikoryu-e30004.html
- Malaysia Geospatial Data Infrastructure (MyGDI)
- UK Geospatial Commission (2019) Guidance: Linked identifier schemes: Best practice guide
- US National Geodetic Survey
- US National Geospatial Data Asset Management Plan
- US FGDC: NSDI | Overview of the National Spatial Data Infrastructure: https://www.fgdc.gov/nsdi/nsdi.html

Principle 2: Geocoded unit record data in data management

2

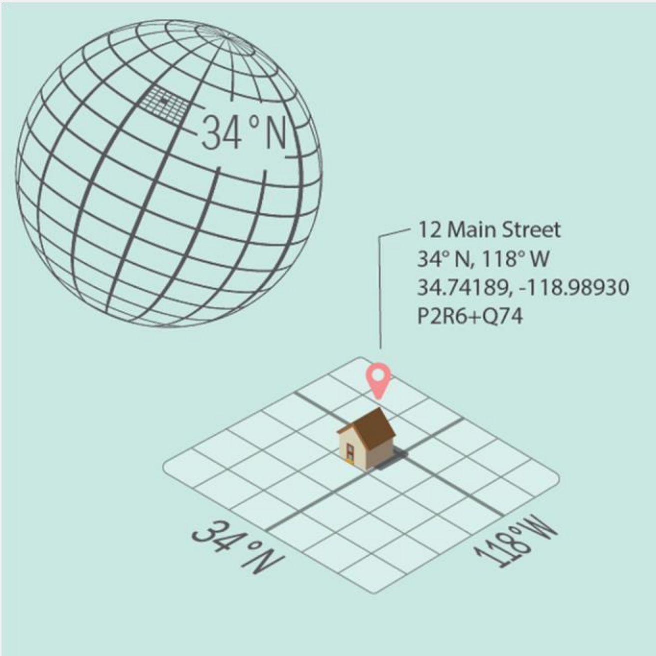
Principle 2 focuses on linking each statistical unit record to a specific geographic location using geocoding or georeferencing with coordinates, small geographic area codes or linked-data identifiers within an NSO's secure data management environment. This allows statistics to be applied to any geographic context, supporting data integration or further data linkage of datasets from other data sources. It also mitigates challenges that arise with new geographies or changes in existing geographies.

? Definition: What does this Principle cover?

The Statistical Commission[18] recommends that all statistical unit record data should be collected or associated with the most accurate location reference possible, preferably a coordinate (i.e., georeferencing). If this is not possible, then the association of the statistical units with the most granular geography (i.e., an area in space, usually called a "polygon") or a grid cell is necessitated, again to the most granular level possible (i.e., geocoding).

Principle 2 relates to the process of linking and storing each microdata/statistical unit record to high-precision geographic references, applying location referencing to the UNGGIM Global fundamental geospatial data themes mentioned in Principle 1.

Example of a geocoded unit



Source: UNGGIM, UNSD (2025)

[18] E/CN.3/2018/33

The use of terminology can vary in geospatial and statistical communities. This publication refers to 'geospatially enabling' data, as all processes of attributing record level data with a location. For additional clarity, we will use the terms:

- "Georeferencing" when referring to the provision of geographic coordinates to record level data.
- "Geocoding" when referring to linking a record to a unique identifier (or linked-data identifiers) whether this is an address, or a geographical place or area (see Principle 3 section). These identifiers might be already attributed with a geographic coordinate which will additionally provide the record level data with a georeferenced.

To ensure data protection, confidentiality and sensitivity, geospatially enabling data should occur within a secure, standards-based environment. Geocoding unit record data allows statistics to be produced in various geographic contexts and facilitate adaptation to changes over time. If precise geocodes or similar location references such as address registers are unavailable, countries should consider alternative more precise point-based georeferencing or even use more general location descriptions. The use of persistent unique identifiers will allow further linkage of non-spatial data, historical information, and support version control of features.

Objectives and Benefits: Why do we need this Principle?

Since 'everything happens somewhere,' location is a common link between many datasets. NSOs gather and share data on social, economic, and environmental topics. These statistical data can be combined with geospatial data from NGIAs to understand and analyze where things happen which could explain some statistical phenomena or understand patterns across space. Principle 2 connects statistical and geospatial data, building on Principle 1, which establishes the system to define locations. Principle 2 then assigns geocodes (and coordinates if possible) to these records, allowing data to be aggregated or disaggregated by geographic areas.

The goal of Principle 2 is to allow for the linkage of all statistical unit records to a precise location, wherever it is possible to do so, and achieve the following objectives:

- **Data Management:** Geographic references available from Principle 1 are embedded in the NSO's digital infrastructures and data management environments (i.e. databases) and centrally managed. This allows for the storage of consistent and interpretable geographic references as a "single point-of–truth", with clear data governance policies with regards to maintenance and custodianship.
- **Data Linkage:** All statistical microdata is 'geospatially enabled' i.e., all statistical unit records should include or be linked to a precise geographic reference (ideally referencing a specific point / coordinate), and if not, the smallest geographic area possible. This permits increased linkage capability and the integration of other datasets from a wide variety of sources using common location attributes.

- **Time Resilience:** Time stamped geospatial data enables longitudinal studies and historical research, despite possible changes in geographies or geographical units through time.
- **Data Disclosure**: can be effectively managed, including the protection of privacy and confidentiality through the aggregation of data at different geographic levels.

Implementing Principle 2, will bring the following benefits[19]:

Strategic Benefits	Benefits for NSOs
• Improved crisis response with greater amount of geospatially enabled datasets supporting better emergency planning and response, enabling quicker and more effective actions during national disasters or public health crises at national level. • Enhanced public services, due to greater availability of geospatially enabled demographic data, identifying unserved areas and optimizing service provision. • Enhanced policy development though increased opportunities for localized/ targeted interventions and resourcing for policy makers, thus reducing the impact on costs. • Greater data integration between increasing numbers of geospatial datasets at national level. • Increased innovation with new ways of looking at social issues and crises. • Provides a platform for feedback and continuous geospatial data enhancement at national level, with NSOs, as users, playing an active role in contributing to improving the quality and currency of geospatial datasets (e.g., addresses).	• Greater understanding of the impact of the physical and built environment on society (e.g., access to amenities, open spaces, etc.). • Enhanced data linkage across data sources, improving accuracy and confidence in matching records (e.g., matching census data with administrative data). • Greater insight at a more granular level based on spatial distribution and patterns (e.g., concentration of records in a specific area which could be investigated further or targeted through localized policies). • Ensures the management of detailed and sensitive data, through the ability to aggregate information and create disclosure control thresholds. • More accurate modeling and therefore predictions, including disaggregation of data.

[19] Answers compiled from a workshop at the Expert Group for the Integration of Statistical and geospatial Information in Nairobi, Kenya in September 2024 and combined with existing published benefits.

Key Stakeholders: Who needs to be involved?

While NSOs typically manage the statistical record georeferencing and geocoding process, maintaining a centralized or federated geospatial data store(s) requires ongoing collaboration with data creators, suppliers, and custodians.

Key stakeholders include:

- NGIAs for geodetic reference frames
- authorities / data custodians for national addressing data
- global and regional institutions (e.g., UNSD, UNGGIM) setting good practices, frameworks for data management
- agencies like ISO and OGC for geospatial data standards
- various NSO units maintaining the base registers
- IT departments for data management support

Requirements: What does this Principle need to be implemented?

It is imperative to have the correct data management tools, techniques, and good practices to facilitate the matching, linking and management of geocodes within statistical datasets.

Governance & Institutional Capacity

It is important to have **clear data governance** establishing defined internal geography data management **policies** at institution level covering:

- The roles and responsibility for data maintenance and custodianship of geographies within or across institutions.
- The rules for data storage, retention and managing changes over time.
- Naming and coding conventions for geographies and addresses.
- The matching and linking process, including clerical matching when automation is not possible.

Human Resources & Capacity

Skilled staff (and the necessary tools) are required to **assign coordinate references** at statistical unit record level. This can be done either through:

- Using mobile apps for data capture, such as digital data collection on the field where records are entered on a mobile device, using its actual localization to attribute location information to the records.
- Methodologies (preferably automated) to match records to the address/geography repositories e.g., using address/ text matching algorithms or geospatial techniques such as reverse geocoding or point in polygon techniques[20].

Data & Interoperability

As described in the above 'Definition' section, data should be collected with the **most accurate location reference** possible, preferably a coordinate (e.g. latitude and longitude). This allows for data aggregation to any geographic level required.

The availability of **addresses** and **geographic boundaries** data is also key for georeferencing and geocoding unit record data so it can be assigned a location.

Technology & Infrastructure

A centralized or federated geospatial data store(s) or database(s), within an effective and **secure** data management environment, will need to exist to **host** consistent and interpretable geographic references:

- For addresses: a managed address register with unique identifiers for addresses or buildings, specific coordinates (centroids) and other relevant information.
- For geographies: a repository for small geographies or grid cells with unique identifiers or codes for each area. This requires the establishment and implementation of data and metadata standards to ensure that geographic references are well documented and consistent across datasets, allowing them to be used effectively over time in different applications.

Access to mobile GIS would be an advantage for field data capture if required.

[20] See Appendices for Definitions of Terms.

Links and Resources

Below are links to further reading or resources relating to Principle 2:

- ISO 19112: *Geographic information – Spatial referencing by geographic identifiers* https://www.iso.org/standard/70742.html
- Georeferencing.org and GeoreferencingQuickReferenceGuide.doc

About georeferencing in statistics:

- UNECE GSBPM and the GeoGSBPM
- GSIM: Introduction to the GISM and https://unece.org/statistics/modernstats/gsim
- European GEOSTAT project - Data management principles
- Guidelines on the Use of Registers and Administrative Data for Population and Housing Censuses

Principle 3: Common geographies for dissemination of statistics

3

Principle 3 focuses on defining a common set of geographies for the dissemination of statistics. This ensures that statistical data is consistent and comparable across different geographic areas and enables aggregation and disaggregation of data to various geographical levels.

Definition: What does this Principle cover?

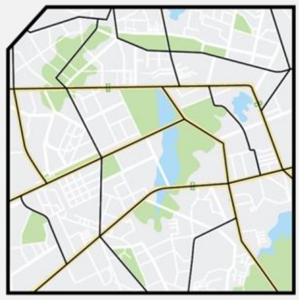
Principle 3 applies geography as a tool for integrating data. It promotes the creation and use of a common and agreed set of geographies for the display, storage, reporting, and analysis of social, economic, and environmental comparisons across statistical datasets from diverse sources, to support informed decision-making.

A common geography is like a "dictionary" of all commonly used geographic areas for a country or region.

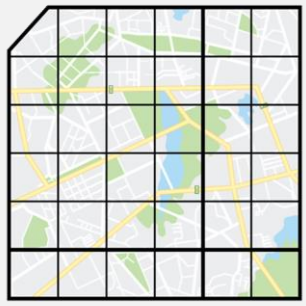
They come in two types (Geography Types box):

- *Administrative geographies* are delimited by polygonal extents (or 'shapes') based on local physical features or other criteria. This is the case for administrative boundaries or statistical units (such as mesh blocks or functional areas[21]) which are defined by political, property, or topological subdivisions.
- *Integrative geographies* are delimited by grids which are defined by the type of coordinate reference systems used and can vary in size depending on the intended use-cases, resolution required and relevant disclosure control rules.

Geography Types



Administrative Geographies



Integrative Geographies

Source: UNGGIM, UNSD (2025)

[21] As per the UNGGIM Global fundamental geospatial data themes - see Appendix 1 on page 30.

Geographic areas define a portion of the Earth surface delimited by physical, human, or environmental characteristics. Geographic areas can vary in size in scope such as territorial, political, natural, economical, educational, institutional and postal. Common geographies can include:

- statistical geographies: statistical units/grid cells, statistical regions, and census areas
- political or administrative boundaries
- functional areas such as economic areas (e.g., workplace catchment, retail, or industrial zones) or recreational and other service delivery areas
- postal areas, usually for mail and other delivery purposes
- cadastral parcels and other land administration area
- integrative[22] geographies, such as national grid-based areas, mesh squares, or Discrete Global Grid systems (DGGS)[23]

Principle 3 establishes the fundamental importance of balancing existing statistical and administrative geographies with other geographic referencing systems, such as grids, and their complex spatial relationships and hierarchies.

Objectives and Benefits: Why do we need this Principle?

Having a common set of geographies enables the production and dissemination of integrated statistics and geospatial information at different geographic output levels within a country to support informed decision-making both at national and local scale. As such, it enables the aggregation of data accurately from the smallest common dissemination geography to higher level dissemination geographies, to meet the needs of the widest possible array of decision makers irrespective of geographic boundaries and/or scale. This includes aggregations at both higher level administrative and statistical geographies (e.g., state, province, or region) and smaller geographies such as electoral areas, census tracts, and other functional areas[24] or supporting the smallest possible cell size when using a gridded geography.

Common geographies enable basic statistical reporting, geostatistical analyses, and dissemination at different scales (such as local, sub-national, national, regional[25] , and global, as well as institutional[26]), where the resulting outputs can be compared and assessed on a

[22] For more information about integrative geographies, please refer to the Definitions of terms in Appendix 2 on page 32.
[23] See Definitions of Terms – Appendix 2 on page 31.
[24] There is nothing inherently incorrect with a unit record-to-small area geography approach. However, this approach will become less optimal over time as the demand from stakeholders for statistical data for new and additional common higher-level geographic areas are made of institutions responsible for integrating statistical data and geospatial infrastructure.
[25] 'Regional' here is used as the UN regions context e.g. Americas, Europe, Africa, Arab states, Asia Pacific, etc.
[26] E.g. Global alliances such as la Francophonie, G20 and others.

consistent basis. These geographies also provide a mechanism to enable the management of privacy and confidential statistical and geospatial data outputs.

The goal of Principle 3 is to support the provision of a common set of geographies that ensure the consistent geospatial aggregation and dissemination of statistical data, irrespective of whether they are in gridded or administrative boundaries, through achieving the following objectives:

- **Consistency and comparability:** using the broadest array of common geographies (legal, administrative, statistical, or integrative) providing a statistical base for data to be integrated and compared across different geographic areas at various scales, supporting geostatistical analyzes and visualization.
- **Geographic hierarchy and currency:** supportive **metadata** on the lifecycle and changes of the geographies through time is required. The **conversion** of data between geographies is supported, through standard conversion mechanisms – usually in the form of correspondence[27] tables for geographic area codes. Conversion between smaller and larger geographies will support aggregation and disaggregation methods, allowing users to flexibly use data at the required level of granularity.
- **Data privacy and confidentiality:** ensuring that country level and international data privacy and quality standards are considered and respected in the design of the common geography sets. Common geographies help manage privacy and confidentiality of statistical and geospatial data by aggregating unit-level data to ensure individuals cannot be identified in statistics.

Implementing Principle 3, will bring the following benefits[28]:

Strategic Benefits	Benefits for NSOs
• Consolidated national core fundamental geospatial datasets through the creation and maintenance of administrative, statistical, and functional areas which stakeholders may use to disseminate their own data. • Flexibility of presentation of statistical data at various geographical scales, across subnational areas supporting localized interventions not just	• Provision of a base for statistical data to be compared across geographical areas. • Uniformed allocation of statistical data to smaller geographies or grid units allowing for aggregation to larger geographies. • Provision of a conversion mechanism for mapping out statistical information between administrative boundaries and/or gridded in space and time.

[27] For more details about "correspondence tables", please refer to the Definitions of Terms section, in Appendix2 on page 32.
[28] Answers compiled from a workshop at the Expert Group for the Integration of Statistical and geospatial Information in Nairobi, Kenya in September 2024 and combined with existing published benefits.

nationwide policies (e.g., for local government).

- Improves resources management efficiency, through identifying and addressing sub-national disparities through various lenses (depending on the geography used).
- Facilitated international comparisons, potentially increasing international cooperation and benchmarking.
- Greater and more effective governance and improved socio-economic outcomes at country level.
- Dissemination of the statistical outputs at various scales depending on user requirements, adapted to disclosure control rules, retaining privacy and confidentiality of input data through data aggregation by geographies.
- Greater opportunity to disseminate data both in a tabular form but also as a map, (e.g., choropleth maps) supporting a wider range of use-cases.

Key Stakeholders: Who needs to be involved?

Key stakeholders in this Principle are similar to Principle 2.

It is also important to promote stakeholder consultation and participation, as per Principle 1, in the creation of new geographies, grids and/or geographic classifications (a classification constituted of one or more common dissemination geographies). This will lead to increased data consistency and efficiency among participating NSOs, NGIAs, international and regional organizations, as well as other relevant institutions. It will contribute to the generation and maintenance of the geographic reference layer, thus feeding back to enriching the "functional areas" from the global fundamental geospatial data theme.

Requirements: What does this Principle need to be implemented?

Governance & Institutional Capacity

At institution level, it is imperative to have clear **policies** and **governance** relating to the maintenance and lifecycle of the various common geographies at national level, defining responsibilities across the various stakeholders. Therefore, the creation of new or bespoke geographies or geographic classifications should be **collaboratively assessed** and acknowledged by interested stakeholders prior to their adoption. They should be integrable within existing and emergent statistical geospatial infrastructures.

Geography policies should cover:

- **Change management** e.g., how changes are applied and implemented, published, etc.
- **Definition, naming and coding methodologies,** documenting decisions when changes occur.

Human Resources & Capacity

At institutional level, there is a strong requirement for skilled staff in **geospatial technology** and **GIS**[29] as well as geospatial **data engineering**. Those skills support a variety of activities relating to the update and maintenance of the common geographies. Specific skills include:

- **Digitizing boundaries** for the geospatial data capture of geographies. Staff should be able to apply **topological checks** to assess and augment data quality. Geospatial boundary files should have no slivers (e.g. gaps) and overlaps.
- **Spatial analysis** skills are also required to understand the hierarchy and the spatial relationships between the various geographies, as well create correspondence tables. These skills can be applied using GIS software or spatial data science (e.g. coding with languages like Python).

Data & Interoperability

There should be an **established inventory** of all administrative, statistical, and other geographies, as well as a **clear understanding of their hierarchical and spatial relationships**. For this, there is a need for:

- A well maintained and consistent framework of geographies at national level consisting of a database of geospatial datasets of each boundary set and associated attributed **names and/or codes for each area**.
- A set of tools enabling the understanding of the spatial relationships between geographies and the processing of geographies – for example, diagrams, correspondence tables, interactive map tools to visualize and query geographies' spatial relationships.
- **Metadata and documentation** for all geographies, including methods for delineation and changes over time, following international standards; information on changes should be

[29] For more information about Geographic information systems, please refer to Appendix 0:

Definitions of Terms

documented using time stamping and lifecycle information on births, splits, mergers, and border shifts within these geographies (such as when new cadastral or political boundaries are created).

Technology & Infrastructure

Geospatial data requires specialist software and technology to support the location elements of the datasets. Deployment of geospatial technology infrastructure can vary (from minimal to extensive) depending on the requirements and budget of the institution involved. It is important to involve information and technology divisions to ensure the geospatial technology is compatible with other technology within the organization. Minimal requirements include:

- **GIS software** to manage, manipulate, analyze and visualize geographies. There are several GIS software solutions available globally, some open source, some proprietary; some desktop-based and some web-based applications. Identifying the right one will depend on institutional circumstances.
- **Spatial tools**:
    - Maintenance and updates to these geographies requires tools to clean and quality assure the data. **Topological tools** can process and validate data based on their topological and spatial relationships.
    - Establishing spatial relationships of datasets using **spatial analysis tools** (calculating distances, whether one feature is within another, etc.).
- **Geospatial databases:** Whilst it is possible to use folders, using a geodatabase is a more effective way to manage and store geospatial data, due to the specificities of geospatial data in terms of format and metadata. There is very good provision of open-source geospatial database (e.g. PostGIS for PostgreSQL or Spatialite for SQLite).

Links and Resources

Below are links to further reading or resources relating to Principle 3:

- OECD: Geographical Definitions
- Indonesia: One Map Policy of Indonesia: Status, Challenges, and Prospects
- European Commission / Eurostat's pan-European geographies: Nomenclature of territorial units for statistics (NUTS) Overview
- US Census Bureau: Geography: Concept and definitions (see Geographic Hierarchy diagram)
- UK Open Geography Portal: Hierarchical representation of UK Geographies (July 2024) and Coding and Naming Policy for UK Statistical Geographies.
- Australia Bureau of Statistics, Australia Statistical geography Standard (ASGS): Correspondences Tables

Principle 4: Statistical and geospatial interoperability in data standards, processes and organizations

4

Principle 4 focuses on making statistical and geospatial data work together seamlessly. It supports increase usability by making data easier to combine across domains through greater interoperability. This improves how we create, find, share, integrate, and use location-based statistics and data such as geospatially enabled statistics and geospatial data.

It will increase the potential application of a larger range of data and technology and make more information available and accessible for decision making. It will facilitate better cooperation between all stakeholders producing and using statistical and geospatial information.

? Definition: What does this Principle cover?

Principle 4 defines the preconditions for statistical and geospatial data to work together in a data ecosystem in which those involved interact with each other to exchange, produce, and consume data. Interoperability between statistical and geospatial data and metadata standards is needed to overcome structural, semantic, and syntactic barriers between data and metadata from different communities and providers.

Data Interoperability refers to the ability of different systems, organizations, and applications to exchange, interpret, and use data in a coordinated manner. This involves ensuring that data can be shared seamlessly across various platforms and understood by all parties involved. Key aspects include data exchange and processing[30] (including business processes); standardization of formats, protocols[31] ; metadata, documentation and data quality[32].

According to the FAIR (Findability, Accessibility, Interoperability, Reproducibility) Principles[33], the "Interoperability" Principle emphasizes metadata. Metadata should be ready to be exchanged, interpreted and combined in a (semi) automated way with other datasets by humans as well as computer systems. This supports the integration between datasets but also the interoperation with applications or workflows for analysis, storage, and processing. It recommends:

[30] What is Data Interoperability? - Reltio
[31] What Is Interoperability and Why Is It Important? | CSA
[32] Data Interoperability | NNLM
[33] Fair principles interoperability section: https://www.go-fair.org/fair-principles/

I1. (Meta)data use a formal, accessible, shared, and broadly applicable language for knowledge representation.

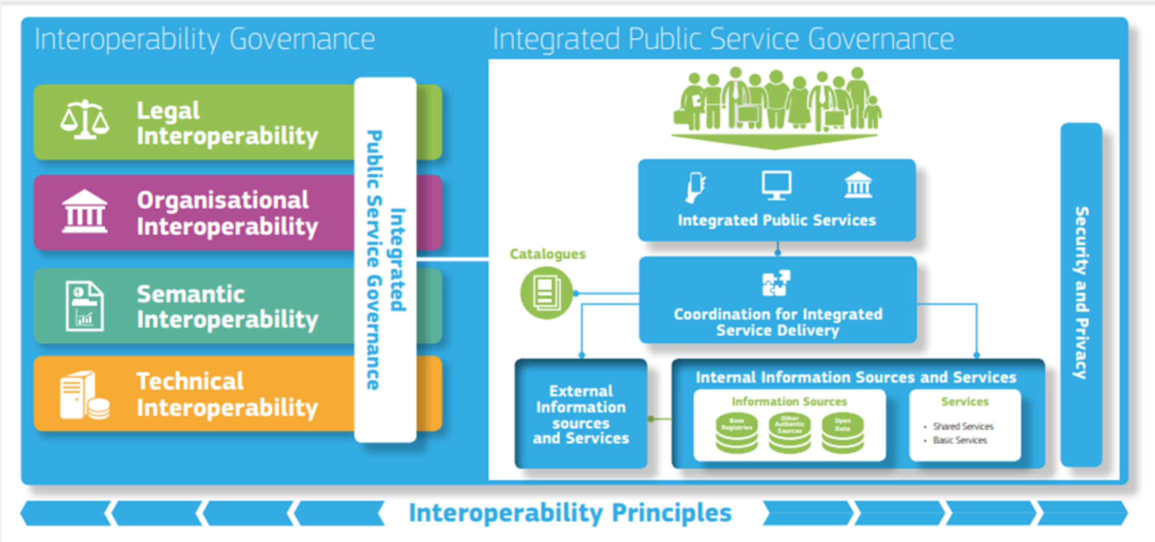
I2. (Meta)data use vocabularies that follow FAIR principles

I3. (Meta)data include qualified references to other (meta)data

Principle 4 extends the interoperability to all data, metadata, standards, and good practices that facilitate the integration and output of geospatially enabled statistical data. This includes tools and methods used in all stages of statistical production. Geospatial-statistical interoperability supports reproducibility, quality management, and the mechanisms by which stakeholders and users interact.

The European Interoperability Framework (EIF) (see box insert) offers a good example of an interoperability model supporting cross nations data sharing. It identifies four interoperability dimensions necessary for effective collaboration and data sharing among public administrations.

The European Interoperability Framework (EIF) Conceptual Model



Source: European Commission, New European interoperability Framework (2017)

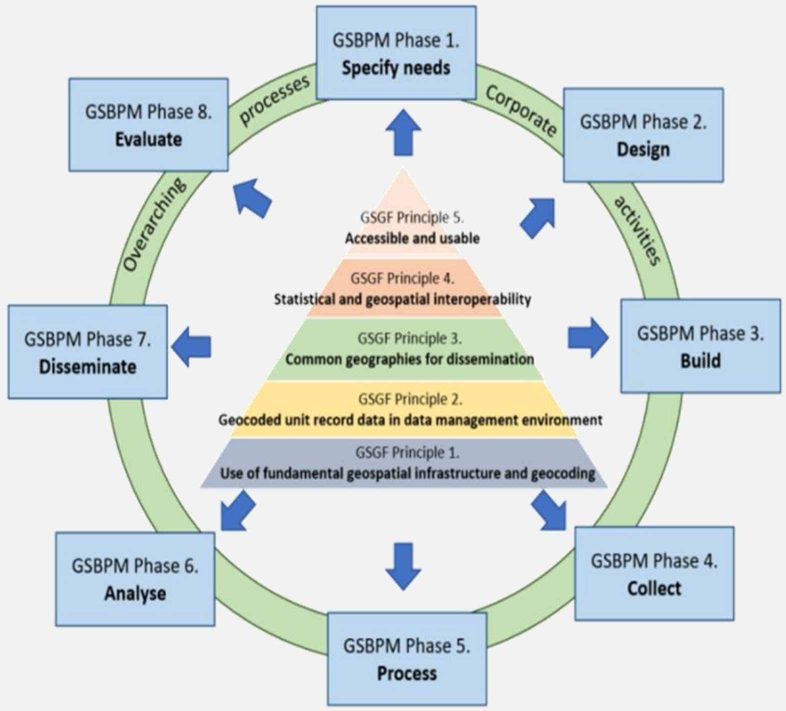
It identifies four dimensions which are crucial for integrating and producing geospatially enabled statistical data and are closely interconnected:

- **Legal Interoperability:** Enables organizations operating under different national legal frameworks, policies, and strategies to work together supported by national laws and policies, which allow cooperation and data sharing.

- **Organizational Interoperability:** refers to the way in which public administrations (i.e., government agencies and organizations) align their business processes, responsibilities, and expectations to achieve commonly agreed goals.
- **Semantic Interoperability:** refers to the exchange of data and information through common terminology and consistent data formats.
- **Technical Interoperability**: refers to systems and services linked through standard interfaces, services, and data formats, including metadata standards.

However, both the statistical and geospatial communities operate their own general data models, metadata capabilities, architectures, and data infrastructures, creating differences in fundamental terminology over time. There needs to be greater incorporation of geospatial processes, standards, and best practices in the statistical business processes and data management systems, not just for sharing and disseminating data.

The Geospatial view of the GSPBM



The diagram of the GeoGSBPM illustrates the GSGF Principles contributing to all phases of the GSBPM.

Considering how to better integrate existing geospatial frameworks, standards, and practices into the Common Statistical Production Architecture (CSPA) will make it easier and more efficient to create, share, find, and use geospatially enabled statistics. This will in turn help combine data from across the geospatial, statistical, earth science, and administrative domains.

The UNECE and ModernStats group have developed the geospatial view of the Generic Statistical Business Process Model (GeoGSBPM)[34]. It seeks to further incorporate geospatial processes within the statistical models, thus increasing interoperability.

The new GeoGSBPM document includes details on geographic information for each phase of the GSBPM, describing activities and roles related to geographic data in its sub-processes.

Objectives and Benefits: Why do we need this Principle?

Greater interoperability between statistical and geospatial data helps overcome barriers between integrating different data sources. This improves the accessibility and usability of data for various applications, enhancing decision-making and cooperation among stakeholders.

Implementation of Principle 4 achieves the following objectives:

- **Efficiency and Simplification:** Improves the creation, discovery, integration, and use of geospatial data and statistics, by developing infrastructure that allows seamless data integration and access.
- **Overcoming Barriers** by addressing structural, semantic, and syntactic difficulties between different data and metadata sources and supporting cooperation through appropriate national legal and policy frameworks.
- **Enhanced Interoperability** by promoting the use of open standards and good practices, to improve the usability of geospatial and statistical data across different applications and systems, including data management systems, data modeling and production planning.
- **Broader Data Application** by expanding the use of various data and technologies, enhancing decision-making through improving the fitness for use of data for various applications and data management systems.
- **Stakeholder Cooperation:** Promotes better collaboration among those producing and using the data through greater flow of data between organizations.

[34] See UNECE GeoGSBPM V1.0, May 2021: https://unece.org/sites/default/files/2024-06/Geospatial%20View%20of%20GSBPM%20v1_0%20Final.pdf

Implementing Principle 4 will bring the following benefits[35]:

Strategic Benefits	Benefits for NSOs
• Greater efficiencies and economies of scale, reducing the need for costly data conversion processes and saving time and resources across the wide range of organizations at national level. • Further opportunities for data sharing between organizations, through seamless data exchange between different systems and organizations. • Improved data availability and data integration of diverse datasets to gain more comprehensive insights. • Greater discoverability and accessibility for users to required information via defined technical and user interfaces that do not require domain knowledge of statistical or geospatial data. • Increased transparency and trust in the data irrespective of the user, as interoperable datasets can be validated in an easier way.	• Reduction of data processing and engineering resources to ingest geospatial datasets (Principle 1) for statistical production. • Reduction in costs and duplication of efforts of data ingest, with tools which can be used, reused, and automated. • Increased speed of data ingestion and usability. • Increased opportunities for reproducibility of statistical data. • Provision of standards and frameworks relating to efficient dissemination of statistical data • Greater integration of data from various sources and institutions such as administrative data. • Reduced risk of mistakes, errors, and data loss during ingestion, leading to higher quality statistical outputs.

Key Stakeholders: Who needs to be involved?

The main stakeholders for setting interoperability standards will be the global and regional standard setting bodies, such as ISO, OGC, IHO and organizations and member states driving the modernization of official statistics (e.g., UNECE[36], OECD).

Whilst some standards are set, agreeing a common approach to implementing these standards will require consultation and cooperation with national stakeholders, which will include:

- NSOs and NGIAs – data custodians, architects, engineers, and managers
- administrative data suppliers (e.g., relevant government institutions / ministries)

[35] Answers compiled from a workshop at the Expert Group for the Integration of Statistical and geospatial Information in Nairobi, Kenya in September 2024 and combined with existing published benefits.

[36] This is not limited to those countries within the geographic bounds of Europe and includes several non-European Member States: https://www.unece.org/oes/nutshell/member_states_representatives.html

- relevant users such as local government (including the emergency services), academia and scientists, third sector (charities), private enterprises, and members of the public
- data suppliers across the broader national data ecosystem

Requirements: What does this Principle need to be implemented?

Governance & Institutional Capacity

The integration of statistics and geospatial information requires **organizational interoperability** – which translates into strong governance and cooperation across different departments within NSOs and NGIAs, as well as input from other government agencies, organizations, and information producers and users. This might include:

- Sharing of data strategies and data policies within the respective custodians' organizations, collaborating in their development with key stakeholders and partner agencies.
- Documenting, integrating, and/or aligning business processes and relevant information, ensuring it meets the requirements of stakeholders (including NSOs).
- Sharing technical issues and collaboratively resolving them.
- Regular reviews to ensure ongoing relevance and effectiveness.
- Adjusting and aligning cooperation models and agreements when necessary.

Policy & Legal

At the national level, **legislation and government policies** should facilitate cooperation and data sharing between stakeholders. For example, national laws and policies should:

- Encourage the publication of open data through government wide or fully open license agreements for non-sensitive information wherever possible.
- Support the development of a national spatial data infrastructure including allowing NSOs (and others) to have access to essential geospatial information at defined quality levels.
- Avoid barriers to cooperation across nations, with clear agreements about how to deal with differences in legislation across borders in some instances, where this occurs.

Human Resources & Capacity

Institutions need to have staff competent in **data architecture** and **data management processes**. Specific knowledge of the intricacies of geospatial data architecture would be an advantage and would help increase interoperability by building common data models and compatible metadata processes considering both statistical and geospatial dataset requirements. Such skills are quite niche, therefore, NSOs should consider **upskilling** data architects in geospatial data literacy and foster a knowledge sharing relationships with NGIAs to learn from their geospatial data specialists.

Data & Interoperability

To overcome the siloed evolution of the statistical and geospatial communities, it is important to build **awareness of the different standards** and terminology used to seek out standards which complement each other or are harmonized. These can be created by various working groups under a governance structure. These working groups can also contribute to improving geospatial data architecture capability within NSOs to enhance interoperability. Appendix 4 provides an overview of the key standards to consider for aligning capabilities between statistical and geospatial communities.

NSOs and NGIAs should therefore work together and seek to align their **data architecture and data management processes** when it comes to geospatially enabled data, following agreed standards in particular:

- Adoption of internationally adopted standards from both communities.
- Required metadata standards.
- Adequate taxonomies, categorizations and tagging for data discoverability.
- Dissemination formats for geospatial data (e.g., file types), using open-source formats[37] when possible, to support greater access.
- Quality and versioning control and methodologies ensuring reliability, timeliness, and relevance of the data.

Technology & Infrastructure

Finally, it is recommended that key organizations host appropriate technical infrastructures which support the use of the **relevant standards** where systems and services are linked through

[37] Some examples of open-source formats for geospatial data are WMS and WFS for boundaries for example, which are API enabled. For more information, please see

standard interfaces, services, and data formats. Aspects include interface and services specifications, and data and metadata standards and formats.

Links and Resources

About Data Interoperability:

- The UNSD (2018): Data Interoperability Guide - UN Statistics Wiki
- UNGGIM, ISO, OGC: UNGGIM Standards Guide: Online Guide to the role of Data Standards in geospatial Information Management
- Go FAIR initiative – "Interoperable" principles.
- FAIR Data Point specification

About Data Architecture and Metadata Standards:

- GSIM: Introduction to the GISM and https://unece.org/statistics/modernstats/gsim
- SDMX: https://unstats.un.org/unsd/methods/itadvgroup/sdmx.htm and https://sdmx.org/
- SDMX V3 Standards Summary of major changes and new functionality (2021): https://sdmx.org/wp-content/uploads/SDMX_3-0-0_Major_Changes_FINAL-1_0.pdf
- Data Documentation Initiative (DDI): https://ddialliance.org/

About Geographic Standards:

- OGC Standards | Geospatial Standards and Resources
- European Commission Knowledge Base: GeoDCAT-AP
- ISO Standards:
  - U.S. Federal Geographic Data Committee – ISO 191** Suite of Standards for Metadata
  - ISO 19109 Geographic Information - *Rules for application schema* https://www.iso.org/standard/59193.html
  - ISO 19115: *Geographic information - Metadata* https://www.iso.org/standard/53798.html
  - ISO 19119 : *Geographic information - Services* https://www.iso.org/standard/59221.html
  - ISO 19123-1: Geographic information - *Schema for coverage geometry and functions* https://www.iso.org/standard/70743.html
  - ISO 19139: *Geographic Information - XML implementation* https://www.iso.org/standard/67253.html

Other resources:

- Esri ArcGIS Blog (2024): Geoenable SDMX with Data Pipelines
- FAO How to Create and Publish Geospatial Metadata: Best Practices for metadata management
- UK Government Guidance: Publish & Manage data: GEMINI and ISO 19139 metadata

Principle 5: Accessible and usable geospatially enabled statistics

5

Principle 5 focuses on making geospatially enabled statistics easy to access, distribute, and use. It helps data custodians release data with confidence, improve its discoverability and accessibility, particularly promoting machine readable web services to provide dynamic data linkage to data. It also supports policymakers in making informed data-driven decisions.

? Definition: What does this Principle cover?

Principle 5 is about the dissemination of geospatially enabled statistics in ways that ensure the widest access and usage possible; releasing data so it can be used for analysis and policy making. Wider use also increases the return on investment on the creation and collection of statistics by enabling the reusability of the data beyond their original purpose.

According to the FAIR principles[38] published in 2016[39], the first step to ensure wide use of data is to have data which is findable and where users "know how [data] can be accessed, (…) including authentication and authorization". The goal of FAIR is also to optimize the re-use of data through clear and well-described data usage licensing to facilitate understanding and use in different settings. Here we define:

- **"Accessibility"** is making geospatially enabled data easily discoverable, available and retrievable. This will require the development and provision of a set of policies, standards, good practices, and technologies necessary to facilitate the availability of geospatial data. This can involve the format in which the data is published and the level of privacy and confidentiality adherence.
- **"Usability"** is about ensuring the data is fit for purpose, trusted by users. Trust in the data equates to high satisfaction rates by users and returning customers. This can be achieved in many ways, but it is important that the datasets are authoritative and can be relied upon in terms of quality[40].

[38] https://www.go-fair.org/fair-principles/

[39] Wilkinson, M., Dumontier, M., Aalbersberg, I. *et al.* The FAIR Guiding Principles for scientific data management and stewardship. *Sci Data* **3**, 160018 (2016). https://doi.org/10.1038/sdata.2016.18

[40] Referring here to the Quality Dimensions (page 22) of the UN Statistical Quality Assurance Framework (2016): https://unstats.un.org/unsd/unsystem/Documents-March2017/UNSystem-2017-3-QAF.pdf

Objectives and Benefits: Why do we need this Principle?

Policymakers need accessible and usable data to make informed decisions. Data custodians must ensure data is released in accordance with laws and policies, while protecting privacy and following good practices.

The goal of Principle 5 is to support the release of geospatially enabled statistical information in a usable and accessible form, in line with our evolving technology, implementing best practice in:

- **Access and Usability:** Ensuring data is easy to access and use for analysis, visualization, and decision-making, using policies, standards (e.g., FAIR principles), and good practices.
- **Data Maintenance:** Maintaining robust data management practices (e.g., quality control, versioning and tracking changes) building trust in the data provided.
- **Privacy and Confidentiality**: Retaining privacy and confidentiality of input data (i.e., people and businesses) whilst producing high quality and informative statistics.
- **Technology and Standards:** Implementing technologies that support linked data, integrated knowledge systems and open formats that adhere to relevant geospatial and statistical metadata standards.
- **Automation and efficiency:** supporting automated data discovery and integration through standardized platforms.

Implementing Principle 5, will bring the following benefits[41]:

Strategic Benefits	Benefits for NSOs
• Increased speed of access for geospatially enabled statistics for downstream data users nationally. • Enabled data dissemination methods supporting greater data usage nationally. • Reduction of data sharing issues and costs across organizations nationally. • Improved availability of datasets for various policy themes, enabling greater analytical insights. • Enabled timely analyses, supporting faster policy and decision making.	• Release of data whilst retaining privacy and confidentiality protocols and reduces risk of data protection breaches. • Enhanced NSO data management practices and quality assurance processes. • Improved data quality of the NSO's statistical datasets and publications. • Increased trust and satisfaction among the NSO's data users. • Greater reach of statistical publications (user / audience) through

[41] Answers compiled from a workshop at the Expert Group for the Integration of Statistical and geospatial Information in Nairobi, Kenya in September 2024 and combined with existing published benefits.

- Maximized opportunities for automated data ingest across government, creating economy of scale for data operational costs through making use of machine-readable web services.
- Automated data linkage and analytical processes using machine-readable linked data methods, creating economies of scale across government.
- Value added to the individual datasets by allowing the faster integration of complementary datasets.

the use of a wider range of platforms and formats provided.

- Increased return on investment on the creation of statistical outputs, by allowing data to be reused for numerous purposes.
- Development of NSO's staff skillsets, through exposure to further technologies.

Key Stakeholders: Who needs to be involved?

Principle 5 emphasizes the crucial role of statistical and geospatial communities in ensuring the safe and effective use of data. It highlights the importance of understanding and engaging a complex and diverse spectrum of stakeholders, relating to the downstream data users of the published geospatially enabled statistical data. Collaboration is also required to establish best practices for safe data release.

Common stakeholders will include (but not exclusively):

- NSO staff, including analysts and researchers
- NGIAs
- relevant government institutions / ministries
- local government and emergency services
- academia and scientists
- NGOs and civil societies (e.g., from charities to neighborhood groups)
- private companies (e.g., retail, insurance, etc.)
- members of the public / private citizens

Requirements: What does this Principle need to be implemented?

Governance & Institutional Capacity

Key institutional policies will need to be in place with regards to:

- Dissemination: Strong internal **disclosure control policies and systems** will need to be in place to understand the appropriate publication thresholds.
- Platform(s) maintenance: Organizational Information Technology (IT) policies must be implemented and enforced through clear governance mechanisms to ensure the resilience and maintenance of the platform(s).

Policy & Legal

NSOs and NGIAs need to be aware of their national **legislation** framework with regards to **data protection and privacy laws**, specifically understanding current constraints and limitations. A wide range of legislative and operational issues can occur when releasing and analyzing sensitive information about people and businesses. Ideally, the national legislation framework should support cooperation and data sharing and, where not sensitive, promote open data publications.

Human Resources & Capacity

NSOs and NGIAs should invest in **capacity building**, regularly maintaining a skilled workforce so staff are able to review and implement all requirements. This includes:

- Knowledge of all aspects of **legality** relating to dissemination, e.g. data protection laws, disclosure control procedures, etc.
- Proficiency in **data management** and **web-based platforms technology**. These skills are necessary to support the maintenance of the platform, as well as the update of the information published. Ideally, staff will have an understanding of **web and user centric design** to ensure the information is organized in a way that is easy retrievable by users (beyond the key words/ metadata and search facilities).

Finally, engaging with data users through vehicles such as focus groups or user surveys can provide distinct benefits for discussing proposals and validating outputs. These efforts can offer an effective mechanism for conducting user research, gathering feedback from diverse stakeholders, and supporting User Acceptance Testing (UAT) of both dissemination platforms and published datasets to enable agile service development and improvement.

Data & Interoperability

NSOs and NGIAs should work together and seek to align their **data architecture and data management processes** when it comes to geospatially enabled data, following agreed standards, including:

- Required metadata standards.
- Adequate taxonomies, categorizations and tagging for data discoverability.
- Dissemination formats for geospatial data (e.g., file types), using open-source formats[42] when possible, to support greater access.
- Quality and versioning control and methodologies ensuring reliability, timeliness, and relevance of the data.

Technology & Infrastructure

Strong technical infrastructure supporting web-based platform(s)[43] are required for the dissemination and publication of datasets in machine-readable formats, using standardized web services and/or linked data methods, maximizing the use of APIs.

The platform(s) used will need to support comprehensive metadata management following international standards for resource discovery: handle data versioning and the **metadata** (see Principle 4) and accessibility standards required and if possible custom visualization, aggregation and eventually data analysis.

[42] Some examples of open-source formats for geospatial data are WMS and WFS for boundaries for example, which are API enabled.
[43] Open sources should be considered for platforms and software wherever possible however it is understood that the decision will depend on the organisations/ countries based on system adequacy, legacy systems, existing skills sets and staff ability, and funding/opportunity.

Links and Resources

Below are links to further reading or resources relating to Principle 5:

About privacy laws, policies, and/or agreed country-level and international privacy protocols:

- United Nations Fundamental Principles of Official Statistics
- Principles and Guidelines for Managing Statistical Confidentiality and Microdata

About standards and interoperability:

- ISO/TC211 Geographic information/Geomatics
- The FAIR Guiding Principles for scientific data management and stewardship
- Go FAIR initiative: https://www.go-fair.org/fair-principles/

About open geospatial data formats:

- OGC - WFS Standard publication (usually for vector datasets)
- OGC - WMS standard publication (usually for raster map images and tiles, but also used for vector datasets)
- GeoJSON: https://geojson.org/
- OGC Features and geometries JSON (JSON-FG): https://github.com/opengeospatial/ogc-feat-geo-json

About geospatial data management for Geoportals:

- Geospatial Commission (2021): Making your data easy to find: Metadata best practice guide for data publishers
- Geospatial Commission (2021) Blog: The secret to great geospatial data portals start with the user
- Geospatial Commission (2020) Research & Analysis: Finding Geospatial data
- Geospatial Commission (2022) Guidance Collection: Best practice guidance and tools for managing geospatial data

About User Research and User Centric Design:

- UK Government, User Research Service Manual: https://www.gov.uk/service-manual/user-research
- UK Government, Design Service Manual: https://www.gov.uk/service-manual/design

About User Engagement:

- U.S. Census Bureau Stakeholder Engagement: Memorandum: 2030 Census Stakeholder Engagement Strategy

Appendices

1. The 14 Global Fundamental Geospatial Data Themes

In 2019, the UNGGIM agreed on a set of fourteen global fundamental geospatial data themes that would provide a themed representation of the world we live in, whether relating to physical or human geographical features. They represent the minimum information required to have a good understanding of our lived-in environment.

The geospatial datasets listed are collected and collated by various agencies, depending on the national context. Frequently, a combination of institutions or government departments will share responsibility of administering the various types of geospatial data available.

The United Nations' 14 global fundamental geospatial data themes.

Source: UNGGIM, UNSD (2019)

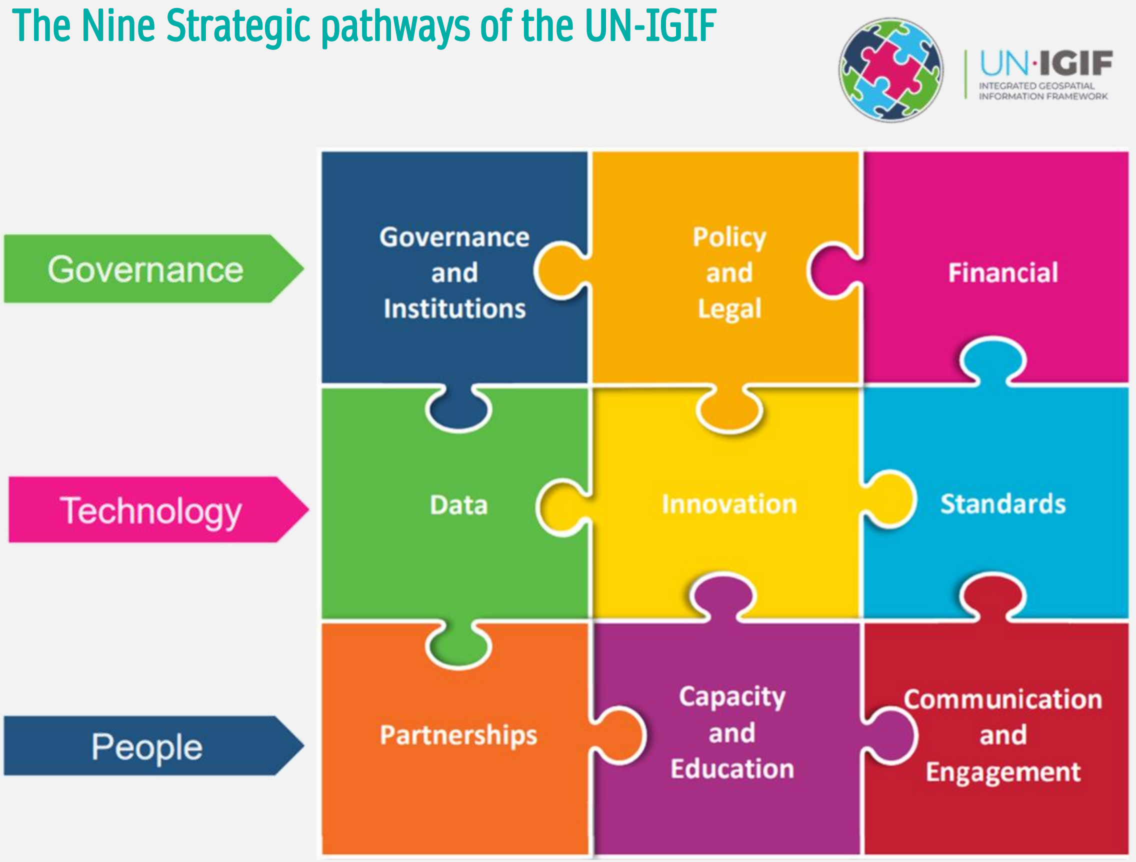
Collation of these datasets may therefore require multiple stakeholder engagement and close coordination between institutions. For example, NGIAs will have a role to play to collect most of the topographical data themes; others such as geological organizations might be responsible for subsurface information datasets.

See UNGGIM Fundamental_Data_Publication.pdf (2019) for more details.

2. The UN-Integrated Geospatial Information Framework (UN-IGIF)

The **UN-IGIF** is a globally developed framework endorsed by the UNGGIM that provides a basis and guide for nations across the world for developing, integrating, strengthening, and maximizing geospatial information management.

The UN-IGIF nine strategic pathways are a tool to group the critical strategic areas where intervention is required to ensure and secure a robust national spatial data infrastructure (NSDI). This encompasses governance, technology, and people and includes: Governance and Institutions, Policy and legal, Financial, Data, Innovation, Standards, Partnerships, Capacity and Education, Communication and Engagement.



Source: UNGGIM, UNSD (2022)

More information about the UN-IGIF can be found on the website: https://ggim.un.org/UN-IGIF/

3. The GeoGSBPM

The GeoGSBPM was developed as an update to the GSBPM, in response to the growing recognition that modern statistical production requires seamless integration of geospatial and statistical information from the earliest stages of planning through to final dissemination. The result identifies specific points within the statistical production workflow where geospatial considerations should be integrated to create more robust statistical products.

By explicitly documenting this guidance, the GeoGSBPM helps statistical organizations understand precisely when and how to incorporate geospatial data within their existing production processes. This structured approach ensures that geospatially enabled statistics achieve higher levels of standardization, interoperability, and analytical capability while maintaining alignment with established statistical production standards.

More information about the GeoGSBPM can be found on the UNECE website:
https://statswiki.unece.org/spaces/GSBPM/pages/312476034/GeoGSBPM

		Phases of the GSBPM							
		Specify Needs	Design	Build	Collect	Process	Analyze	Disseminate	Overarching Processes / Corporate Activities
Principles of the GSGF	1	When assessing data availability, the existence and availability of suitable geospatial information should be first identified from authoritative sources within the National Spatial Data Infrastructure (NSDI)	Geospatial variables (geographies) should be designed for the statistical level unit. Using point-based location as the base geospatial variable will provide considerable adaptability to changes overtime and flexibility to aggregate up to various dissemination level geographies.		Geocoding should be conducted for each statistical unit that is collected and at the most detailed level (e.g. point-based geocoding as opposed to area -based geocoding)	Standardization should take place before the integration of datasets. It can be done through, for example, matching location information in the datasets with centralized standard systems (e.g. address matching, geocoding) which should be based on the national geospatial information context.			Quality Management include: Identify the authoritative (external or interna) sources of reference data and establish quality profile of reference data
	2		The design of components includes: point of entry validation for geographical information; matching strategy; and, spatial analysis.			The mechanism of matching or geocoding the statistical unit record established in Design phase should be consistently applied.			Quality management include: Develop quality dimensions and metrics to be used at different stages and a consistent matching strategy

Principles of the GSGF	Phases of the GSBPM: Specify Needs	Design	Build	Collect	Process	Analyze	Disseminate	Overarching Processes / Corporate Activities
3	Needs of users in terms of geographies (size of unit, type) is discussed. Implications (e.g. cost, reliability, quality) should be communicated and consulted with users	When grid geographies are used, the choice of grid system should take existing regional and global systems into consideration		Inaccuracies in geospatial information detected during field collection should be documented and transferred to the central geospatial information system for maintenance and update (if permitted under statistical confidentiality rules)				
4		Design of all production components should take into account standards used in the geospatial community.	Geospatial services have a broad stakeholder group, statistical organisations should check and consult with service inventories of stakeholders before building components on their own.			When preparing the analysis output, it is important to pay attention to the semantic interoperability so that the output can be used without ambiguities by the users from different domains	International standards should be used as a norm to ensure the products can be found and consumed easily across a range of various user groups from the public and private sectors	Alignment of harmonization of geospatial metadata concepts with those of statistical metadata is critical
5	Discussion of the output format is useful as users for high spatial resolution data (e.g. city, municipal authority) might require data to be provided in certain formats that are digestible with their GIS system. Implications of the size of the geographic units in terms of confidentiality risk should be discussed with users	Design of these outputs should also take potential downstream uses into consideration. Accessibility and usability of geospatially enables statisticas and services can greatly increase by use of standards and open data formats.	Metadata elements are put together during development of dissemination compenents so that they can be disseminated along with the data products and services. To make it more findable and accessible for both internal and extarermal users, metdaat should be documented using standard taxonomy and vocabulary.			Cataloguing and tagging the content using relevant metadata standards can greatly increase the usability of the analysis outputs. Geospatial products components should be cross-checked with other components (e.g. tabular aggregates) before release so that they do no breach privacy on their own as well as in combinations with other outputs.		Statistical organisations are encouraged to explore the semantic web standards as a long term strategic objective with successive milestones to achieve dissemination of data and metadata within the framework of Linked open data (LOO)

Source: UNECE (2021).

4. The GSGF Five Principles and Key Elements Matrix

The table below provides a summary of the inter-dependencies between the GSGF Five Principles and Key Elements.

PRINCIPLES		Governance & Institutional Capacity	Policy & Legal	Human Resource & Capacity	Data & Interoperability	Technology & Infrastructure
	1	National Agency in charge of geodetics; Custodianship of geospatial data themes clearly defined.	National Legal framework or Data sharing agreements, MoUs, etc. to support greater data exchange and collaboration.	Geospatial data literacy across the government organizations; a skilled GIS team in NSOs.	Datasets which are authoritative and in line with the UNGGIM global fundamental geospatial datasets.	The technical infrastructure at national level to host the NSDI and NSS.
	2	Data Governance and Data Management Policies.	N/A	Skilled staff are assigning coordinates and geospatial references (spatial techniques/ code).	Address Register Geographic boundaries repository.	Data stores and geospatial databases handling geospatial reference and metadata; mobile data capture apps.
	3	Policies about maintenance and lifecycle of common geographies e.g. change management, geography naming and coding frameworks.	N/A	Skilled staff in GIS, maintaining and updating the geographies, using topological tools and spatial analysis; data engineering knowledge.	Sets of geographies with names and codes; Data files supporting the understanding of spatial relationships (e.g. correspondence tables).	Geospatial tools (GIS or code) performing topology / spatial analysis; database for storage and geospatial software for viewing the geographies.
	4	Strong governance and organization interoperability: including sharing data strategies and documenting business processes.	Legislation and government policies supporting data sharing between stakeholders; open data policies.	Skilled staff in data architecture and data management processes (including metadata).	Awareness and Collaboration on data standards and interoperability standards.	Platforms supporting the use of agreed metadata and interoperability standards.
	5	Strong disclosure control policies and business processes.	National data protection and privacy laws.	Skilled staff such as Geospatial experts, Developers; Focus Groups feedback on outputs.	Alignment of data architecture and data management processes for geospatial data.	Strong technical infrastructure to disseminate data (e.g. web-based platforms; use of APIs).

5. Key Data Architecture and Metadata Standards for Interoperability

Geospatial data has often been developed independently of statistical organizations. Because of this separate development, there are differences in fundamental terminology between geospatial information and statistical information. The statistical community and geospatial community have therefore operated under distinct sets of standards.

N.B: It is important to recognize that there are many standards, and this appendix does not aim to provide an exhaustive list. This is an attempt to highlight the most relevant standards in use at the time of publication.

The **statistical community** operates on a set of standardized approaches:

- **Generic Statistical Information Model** (GSIM)[44][45] , developed by United Nations HLG-MOS: the most common framework designed to standardize and improve the management of statistical information through a common structure describing, organizing, and linking statistical data and metadata and processes, ensuring consistency and interoperability within the statistical domains.
- The **Standard Data and Metadata Exchange** (SDMX)[46] is an ISO standard (**ISO 17369**[47]) designed to describe statistical data and metadata, normalize their exchange, and improve data sharing. SDMX 3.0[48] also explicitly supports geospatial data.
- **Data Catalog Vocabulary** (DCAT)[49] used to describe datasets in data catalogues and provide a standardized vocabulary to increase discoverability.

The **geospatial community** uses a breadth of standards for data concepts and modeling, for terminologies, georeferencing, data collection, data content, data access, metadata, etc.

- The **General Feature Model** (GFM) is a model of the concepts required to classify a view of the real world – it refers to ISO 19109[50] *Rules for application schema* and coverage geometry in ISO 19123-1 *Schema for coverage geometry and functions.*
- The **ISO 19115**[51] metadata standard broadly provides a conceptual schema on metadata presented as UML diagrams. It provides information about the identification, the extent, the quality, the spatial and temporal aspects, the content, the spatial reference, the portrayal, distribution, and other properties of digital geographic data and services. It has

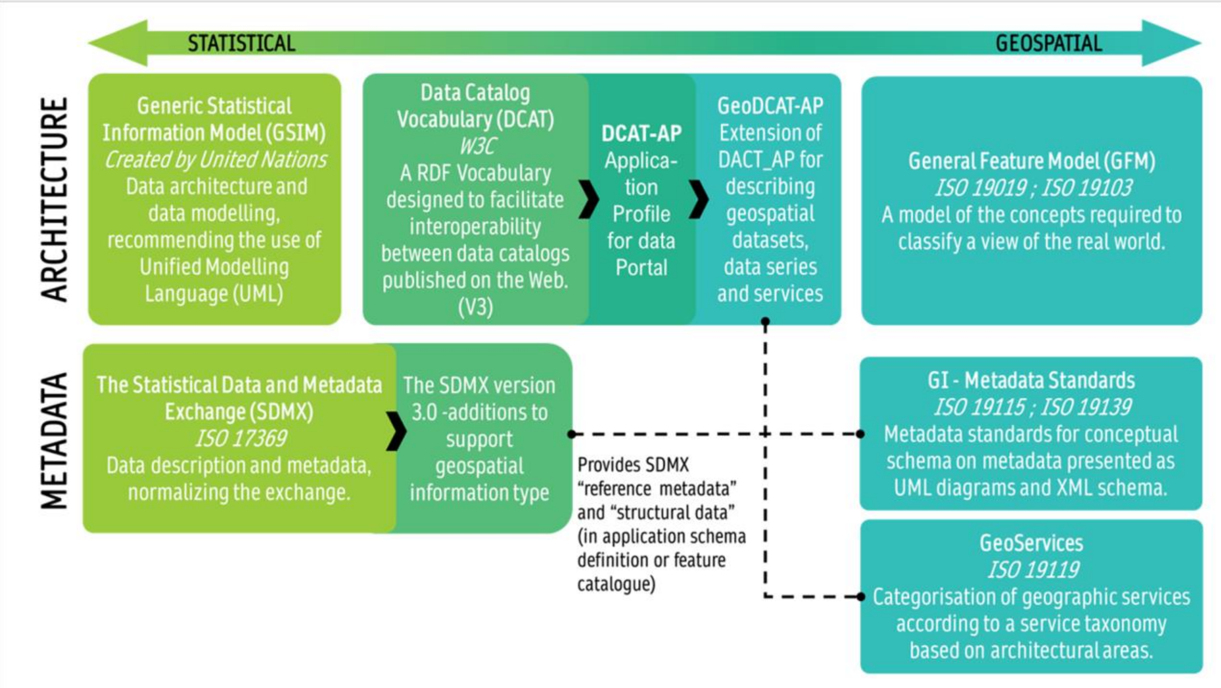
[44] UNSD (2015), Generic Statistical Information Model (GSIM): Statistical Classifications Model
[45] UNECE Statswiki Generic Statistical Information Model page: https://statswiki.unece.org/spaces/gsim/overview
[46] https://sdmx.org/
[47] https://www.iso.org/standard/52500.html
[48] https://sdmx.org/wp-content/uploads/SDMX_3-0-0_Major_Changes_FINAL-1_0.pdf
[49] W3 Data Catalog Vocabulary (DCAT)V3: https://www.w3.org/TR/vocab-dcat-3/
[50] https://www.iso.org/standard/59193.html
[51] https://www.iso.org/standard/53798.html

been implemented widely across the globe, sometimes under various names e.g., INSPIRE in Europe. It has various sub standards for the data types of vectors (vs imagery), plus several application specific standards and good practices to support interoperability of data.

- **ISO 19139**[52] defines Geographic Metadata XML (geographic metadata) encoding, an XML Schema implementation derived from ISO 19115. The UK for example, implemented the Gemini[53] UK standard for spatial metadata – based on the ISO 19139 and considering the needs for INSPIRE too (i.e., ISO 19115).
- **ISO 19119**[54] defines how to categorize geographic services according to a service taxonomy based on architectural areas.

The diagram below highlights the key standards used in both statistical and geospatial communities which enable compatibility (at the time of publication).

Statistical and Geospatial data architecture and metadata standards



Source: UNGGIM, UNSD (2025)

[52] https://www.iso.org/standard/67253.html
[53] https://guidance.data.gov.uk/publish_and_manage_data/harvest_or_add_data/harvest_data/gemini/
[54] https://www.iso.org/standard/59221.html

6. Definitions of Terms

Administrative Geographies	Spatial units legally defined by governmental or administrative boundaries used for governance and public administration. These include countries, states/provinces, counties, municipalities, and other legally defined territorial units. While officially recognized and used in government operations, their boundaries may change over time due to political decisions or administrative reforms. In statistical contexts, these geographies serve as frameworks for data collection, aggregation, and dissemination, aligning statistics with policy-making jurisdictions.
Authoritative Data	Data officially maintained and provided by a recognized authority as the most reliable and accurate source for the specific information, typically maintained by a designated government agency or organization with legal mandate or recognized expertise in that domain. This data serves as the definitive reference for decision-making and ideally adheres to documented quality standards, underdoes regular maintenance, and is subject to stringent governance and oversight. For more reflection on the definition of "authoritative", please refer to the UNGGIM E/C.20/2023/16/Add.2 paper titled Authoritative Data in an Evolving Geospatial Landscape presented at the UNGGIM thirteenth session by the Working Group on Policy and Legal Framework (2023).
Big Data	The topic of Big Data includes datasets that are challenging to use whether due to high frequency of collection, large number of records, and/or heterogenous observations, as well as the data science ecosystems and techniques that produce, analyze, and integrate these datasets into Official Statistics and other products. The opportunities afforded by Big Data include generating novel sources for collecting or imputing Official Statistics, developing, and tracking more sensitive indicators for Sustainable Development Goals, and facilitating high-level integration across disciplines to address complex issues such as coordinating disaster response. The incorporation of location data into Big Datasets allows them to be recruited into geospatially enabled workflows and organizes them in ways that are meaningful to end users and policy makers. Please refer to Annex B for a discussion on the considerations in creating and using geospatially enable Big Data, produced in collaboration with the UN Committee of Experts on Big Data and Data Science for Official Statistics.
Common Geography	A common geography is an agreed set of geographies for the display, storage, reporting, and analysis of social, economic, and environmental comparisons across statistical datasets from diverse sources. Common

	geographies enable the production and dissemination of integrated statistics and geospatial information within a country or region to support informed decision-making. Ideally these national common geographies will align with global geographies to facilitate cross-national comparisons.
Correspondence tables	A correspondence table, also known as conversion or geographic lookup tables, is a tool for the linking of statistical units, used for classifications and for geographic areas. It is designed to help users to mathematically convert statistical data to and from geographic regions or convert between geographies (either current or old to new).
Data Management Environment	A data management environment holistically encompasses the tools, storage, and environment for acquiring, validating, storing, protecting, and processing required data to ensure the accessibility, reliability, and timeliness of the data for its users.
Discrete Global Grid System (DGGS)	A DGGS is a spatial reference system that tessellates the Earth's surface into a hierarchical series of discrete, uniquely identified cells at progressively finer resolutions. Unlike traditional coordinate reference systems based on ellipsoidal models and map projections, DGGS are designed as frameworks for information distinct from conventional coordinate reference systems originally designed for navigation. All DGGS level-0 cells are equal area faces of regular polyhedra, with the most common choices being Platonic solids—tetrahedron, cube, octahedron, dodecahedron, and icosahedron. The geodetic foundation may use simplified geoid models, reference spheres, or non-geodesic surfaces (such as perfectly spherical surfaces) rather than traditional ellipsoidal datums. By directly discretizing the Earth's surface rather than a planar cartographic projection surface, DGGS represents the Earth as a hierarchy of equal area cells with progressively finer geospatial resolution. Individual observations can be assigned to a cell corresponding to both the position and size (or uncertainty) of the phenomenon being observed. DGGS are also called geodetic grids. Officially adopted as an OGC standard in 2017, DGGS enable efficient integration and analysis of multi-resolution geospatial data through their polyhedron-based equal-area cell structure. More information on the DGGS: visit http://www.opengeospatial.org/projects/groups/dggsswg
Enumeration geographies	Enumeration geographies are the smallest geographic units used for census data collection, designed to be assigned to individual census staff. These areas use clearly visible landmarks (roads, rivers, buildings) as boundaries to help field workers easily identify their assigned territory. They serve as building blocks that can be grouped into larger enumeration districts for operational management and later aggregated for data analysis.

Fundamental Geospatial Infrastructure	A Fundamental Geospatial Infrastructure encapsulates the NSDI, standards, technologies, policies, and best practices, amongst other key elements to enable the provision of geospatial information within a country. The UNGGIM endorsed United Nations Integrated Geospatial Information Framework (UN-IGIF)[55]can facilitate the establishment of this infrastructure if such country-level infrastructure is not yet established.
Geocode	Geocodes are, preferably, fine scale geospatially referenced objects that are stored as a geometry data type, such as location coordinates (such as latitude/longitude, or x, y, z coordinates), and/or small area geographies (e.g. mesh blocks, block faces or similar small building block geographies). Larger geographic units, such as enumeration geographies, can be used as geocodes where finer scale geospatial units are not available. The linkage of a geocode to a statistical unit record can occur through use of standard geographic coding systems, a Uniform Resource Identifier (URI) or through other computer-based linkage mechanisms. See "geocoding"
Geocoding	The Global Statistical Business Process Model (GSBPM) defines geocoding as "assigning codes related to geographical places," however, in geospatial terms, it is the process of attributing a record to a specific location (address or coordinate references, not just a code) relevant to the record. This should reflect where the data was collected, where a process occurred, or For the purposes of the GSGF, geocoding is defined as the process of geospatially enabling statistical unit records so that they can be used in geospatial analysis. More specifically, geocoding is the process of linking unreferenced location information (e.g., an address), that is associated with a statistical unit, to a geocode (i.e., a geospatially referenced object). Alternatively, the geocode can be directly incorporated into the statistical unit record. See "geocodes"
Geodesy / Geodetic	Geodesy is the science (branch of mathematics) dealing with measuring the shape and area of the earth, its movement and gravity. For the purposes of the GSGF, it is the science that support the geo-referencing of physical features on earth.

[55] http://ggim.un.org/meetings/GGIM-committee/8th-Session/documents/Part%201-IGIF-Overarching-Strategic-Framework-24July2018.pdf

Geographic Feature	A geometric representation of a feature. This can be a physical feature such as a unit record, a dwelling, or property or a functional area such as an administrative boundary or an economic area.
Geographic Information Systems	A Geographic Information System (GIS) is a system that integrates computer hardware, software, and geographic data to capture, manage, analyze, and display all forms of geographically referenced information. Often referred as digital mapping, it offers more than just cartographic capability, specifically the ability to analyze data based on spatial relationships.
Georeferencing	Georeferencing is a set of broad processes that includes geocoding. It is the process of referencing data against a known geospatial coordinate system, by matching to known points of reference in the coordinate system (e.g. image rectification to survey points or addresses linked to parcel centroids), so that the data can be viewed, processed, queried and analyzed with other geographic data.
Geospatially Enabled Statistics	Statistical outputs that have been processed geospatially (i.e. every statistical unit record data that are linked to location references, such as geospatial coordinates referencing a point, or an address or smallest geography possible if a point reference is not available), making the outputs usable in a geospatial data and analytical environment, and compatible and interoperable with other geospatial datasets.
Integrative Geographies	Geographies designed to efficiently integrate social, economic, and environmental data, usually through a grid-based approach. In Europe, the GEOSTAT population-grid dataset is promoted as a first example of a European Union (EU) population grid.
Interoperability	Interoperability is defined as the ability of systems and data to be universally exchangeable, accessible, reusable, and comprehensible by all parties, enabled through the application of open standards.
Location information	Location information relates to records of what we do and where we do it, thus data which describes objects events or other features and which is spatially referenced to a specific location on Earth. It can be referenced by coordinates (based on a geographic coordinate reference system) or by other location descriptors such as addresses, property or building identifiers, enumeration geographies and other standardized and non-standardized textual descriptions of a location (e.g. village names).

National Spatial Data Infrastructure (NSDI)	National Spatial Data Infrastructure is the framework, technology, policies, standards, good practices, and human resources necessary to acquire, process, store, distribute, and improve utilization of geospatial data.
Reverse geocoding	See “geocoding”
Statistical Geographies	Geographies defined by a set of rules or a methodology meant to create a geographic concept based on population or statistical characteristics (e.g., the metropolitan regions or core-based functional areas, labor market areas outside of metropolitan regions).
Statistical Unit Records	Statistical Unit Records can include persons, households and living quarters, businesses, buildings, or parcels/units of land. Also known as micro data in statistics.
Uniform Resource Identifiers (URI)	Published in January 2005, the specification ‘RFC 3986’ defines the syntax and semantics of Uniform Resource Identifiers (URIs), providing a standardized method of identifying unique resources on the internet. URIs are composed of a unique sequence of characters and can identify virtual resources such as webpages or digital repositories, or physical resources such as books or people. URIs encompass both URLs (which locate resources) and URNs (which name resources).
Unique Identifiers (UIDs)	Unique identifiers (UIDs) are strings of characters used to uniquely identify a resource, such as a record, object, or entity, within a system or database. They ensure that each item is distinguishable from all others, preventing ambiguity and enabling efficient data management. They serve as primary keys that ensure no two items share the same identifier. This allows for joins and merges across datasets, avoiding duplication and facilitating data integrity.

Credits

- Icons were sourced from Microsoft 365 stock icons or Flaticon.com within the free downloads allowance - specifics credits to: Freepik, Kiranshastry, meaicon, awicon and dreamstale.
- Images sourced via Unsplash.com or Adobe stock images.

UN - GGIM
UNITED NATIONS
COMMITTEE OF EXPERTS ON
GLOBAL GEOSPATIAL
INFORMATION MANAGEMENT

Statistical Commission

United Nations
New York, 2025