GSGF Europe

GEOSTAT 4

Title: GSGF Europe

Author: GEOSTAT 4

Project: Eurostat ESSnet grant project GEOSTAT 4

Grant agreement number: 945503 - 2019-FI-GEOSTAT4

Executive Summary

Europe is confronted by a mix of societal, economic and environmental challenges such as climate change, ageing societies and massive immigration from developing countries. Understanding these major cross-border challenges and taking the right decisions requires new insight that we can only derive from new types of data and their combination. Through the adoption by the United Nations Committee of Experts on Global Geospatial Information Management (UN-GGIM) in 2018 and the United Nations Statistical Commission (UNSC) in 2019, the Global Statistical Geospatial Framework (GSGF) has been recognised as a framework for the world that provides an underlying mechanism to integrate statistical and geospatial information.

As GSGF understands it, the integration means real collaboration and knowledge exchange between the geospatial and statistical domains. The goal is that statistics, as well as all other data producers, use the same sources of geospatial information and the same geocoding services - leading to more efficient use of resources and improved integration of data produced by different public organisations and the private sector.

The GSGF facilitates the production of harmonised and standardised geospatially enabled statistical data. It does this by defining five principles which aim to

1. develop organised geospatial infrastructure
2. enable geospatial consistency in data
3. ensure common geographies
4. enable interoperable data
5. provide easy access and usability to geospatially enabled data

Furthermore, the GSGF recognises four key elements that have an important enabling role for the application of GSGF principles:

- standards and good practices
- technical infrastructure
- laws and policies
- institutional collaboration

If GSGF in Europe is to be successfully integrated, it is essential to understand how these key elements are interpreted and how the statistical geospatial community is structured and operates – and how GSGF principles are linked to this European operating environment. To ensure the statistical and geospatial communities can share the same view of their operating environment, and that they discuss the same concepts on common ground, the first steps to describe the GSGF Europe Reference Architecture have been taken.

Frameworks are not developed and used in isolation, but must be seen in the context of an entire suite of other frameworks. Several other frameworks are applied for the purposes of statistical and geospatial domains in Europe, and the GSGF also links to these. Furthermore, four highly current topics for future enhancement are recognised. These are quality, data collection, confidentiality and innovation.

The implementation of the GSGF is very much based on national capability building, cooperation, inspiration and the motivation of the two communities for a common goal to monitor and align business processes - from intra-institutional silos to inter-institutional dataflows and services. This document is intended as a high-level summary of the conceptualisation and interpretation of the GSGF in the European context. It also aims to serve as a general introduction to the GSGF for actors engaged in activities at other organisational levels. It does not replace the descriptions and guidance provided at the global level, but moves towards the framework's regional implementation.

GeoStat

Content

1. List of Abbreviations

API – Application Programming Interface

CoP – Code of Practice

CSDA – Common Statistical Data Architecture

CSPA – Common Statistical Production Architecture

DG – Directorate-General

DGGS – Discrete Global Grid System

DWG – Domain Working Group

EA – Enterprise Architecture

EEA – European Environment Agency

EFGS – European Forum for Geography and Statistics

EG–ISGI – Expert Group on Integration of Statistical and Geospatial Information

EIF – European Interoperability Framework

ESS – European Statistical System

EU – European Union

FAIR – Findable, Accessible, Interoperable, and Reusable

GAMSO – Generic Activity Model for Statistical Organisations

GeoGSBPM – Geospatial view of Generic Statistical Business Process Model

GIS – Geographic Information System

GISCO – ESS working group on Integration of Statistical and Geospatial Information

GSBPM – Generic Statistical Business Process Model

GFGS – Global Forum for Geography and Statistics

GSGF – Global Statistical Geospatial Framework

GSIM – Generic Statistical Information Model

HLG-MOS – High-level Group for the Modernisation of Statistical Production and Services

HVD – High-Value Dataset

IGIF – Integrated Geospatial Information Framework

INSPIRE – Infrastructure for Spatial Information in the European Community

ISO – International Organization for Standardization

IT – Information Technology

JRC – EU Commission Joint Research Centre

MOU – Memorandum of Understanding

NFGS – Nordic Forum for Geography and Statistics

NGIA – National Geospatial Information Agency

(N)SDI – (National) Spatial Data Infrastructure

NSI – National Statistical Institute

OGC – Open Geospatial Consortium

OECD – Organisation for Economic Co-operation and Development

QAF – Quality Assessment Framework

SDG – Sustainable Development Goal

TOGAF – The Open Group Architecture Framework

UN – United Nations

UNECE – United Nations Economic Commission for Europe

UN-GGIM – United Nations Committee of Experts on Global Geospatial Information Management

UNSC – United Nations Statistical Commission

2. Introduction

The implementation and monitoring of European and global sustainable development programmes (UN Sustainable Development, European Green Deal) require reliable and relevant information. The EU's Open Data Directive aims to strengthen the data economy by increasing the amount of publicly held and publicly funded data available for reuse. The use of data from a range of sources and types for multiple purposes requires integrations with a common reference system of harmonised concepts and methodologies, as well as integrations with a common location and temporal framework. Statistical and geospatial communities have therefore identified a common task for building frameworks that support the production of more relevant, detailed and timely information.

Geospatial data plays a key role in revealing insights, patterns and trends, and thereby supports decision making. The adoption of the Global Statistical Geospatial Framework, the GSGF (UNSC and UN-GGIM, 2019) is one acknowledged and important step towards more coordinated geospatial information practices, and the better global integration of statistical and geospatial information. The focus of the GSGF is on the comparability of statistical outputs, preferably in smaller geographical areas, harmonised geospatial data sources, common methodologies and the interoperability of data and metadata to produce efficiently geospatial statistics.

Geospatial and statistical data governance is crucial for avoiding the duplication of work and enabling the sharing of data. Users of statistical and geospatial information do not want to worry about preparing the data before integrating and using it. It should be fit for purpose, ready to use, available in harmonised formats, easily accessible, and have the correct spatial and temporal resolutions to suit the analytical question or pending decision (Eurostat, 2019). The aim of the GSGF is to enable the integration, comparison and sharing of data and utilisation of new sources of data at local, sub-national, national, regional and global levels. The GSGF thereby supports informed, data-driven and evidence-based decision making within and between countries and thematic domains.

To support this work, this GSGF Europe document aims to adapt appropriately the global GSGF to the European statistical and geospatial operating environment. Preferably, this will also have further influence and impacts on the overall data business ecosystem. This document is intended as a high-level summary of the conceptualisation and interpretation of the GSGF in the European context. In addition, it aims to serve as a general introduction to the GSGF for actors engaged in activities at other organisational levels. It does not replace the descriptions and guidance provided at the global level, but rather builds a fundament for the regional implementation of the framework in Europe.

The development of the global GSGF in parallel accomplished the first three GEOSTAT projects (Annex III), leading to the GEOSTAT 4 project and the realisation of this GSGF Europe document. This document presents the culmination of a decade of joint European efforts to improve statistical geospatial integration. Like the descriptions and guidance provided for the global level, this document does not itself intend to provide a detailed plan or design for its implementation. Instead, it allows a considerable degree of flexibility for action. Furthermore, this document does not override the need for, but rather supports, countries in defining their own guidelines for implementation at a national level. This document, along with other available material to support the implementation of the GSGF in the European context, is available on the website of the EFGS.

To realise the potential value of statistical, geospatial and administrative data, a common understanding and working arrangements between communities are needed. In this regard, this document targets three major communities, or groups of stakeholders and their different roles, in terms of data provision and data integration:

- Statistical community
  - NSIs or other public institutions responsible for the production of official statistics
  - European institutions responsible for the governance and coordination of the European Statistical System (Eurostat)
  - European actors responsible for statistical collaboration and cooperation
  - Global bodies and associations responsible for statistical collaboration and cooperation (e.g. UNSC)
- Geospatial community
  - National geospatial agencies and other public institutions responsible for the production of authoritative geospatial data
  - National geospatial agencies or other public institutions responsible for the governance and coordination of National Spatial Data Infrastructures
  - European institutions involved in the implementation and monitoring of the INSPIRE directive (European Commission, JRC, EEA, etc.)
  - European actors responsible for geospatial collaboration and cooperation (UN-GGIM: Europe, EuroGeographics)
  - Global bodies and associations responsible for geospatial collaboration and cooperation (e.g. UN-GGIM, OGC)
- Administrative and other data communities
  - National public institutions responsible for administrative data collection and maintenance of public administrative data repositories (population registers, land registers, tax authorities, business registers etc.)
  - European institutions responsible for legal frameworks for administrative data, e.g. the Open Data Directive (European Commission)

- Businesses, data producers, analysts and users interacting and collaborating with the statistical and geospatial community in the context of geospatial information and gaining benefit in using this framework. Administrative and other data providers are becoming increasingly important as more and more statistics are based on these data sources.

This document, together with the global GSGF guidance material, aims to support communities

- to sketch their current and preferred future states,
- to analyse the challenge of geospatial data from different aspects,
- to identify required actions and priorities,
- to take steps towards a more integrative data environment via the harmonisation and coordination of operations, and
- to work together towards the common objectives in the integration of statistical and geospatial information in Europe.

This document has three main parts. First, the GSGF and its five principles are interpreted for this GSGF Europe document. Second, the GSGF and its five principles are assessed from different aspects which should be considered to achieve a meaningful statistical geospatial framework. These aspects include the basic structure and operation of European statistical geospatial production and the four key elements playing a critical enabling role in the implementation of the GSGF in Europe. Four topics for future enhancement are also discussed. These are quality, data collection, confidentiality and innovation.

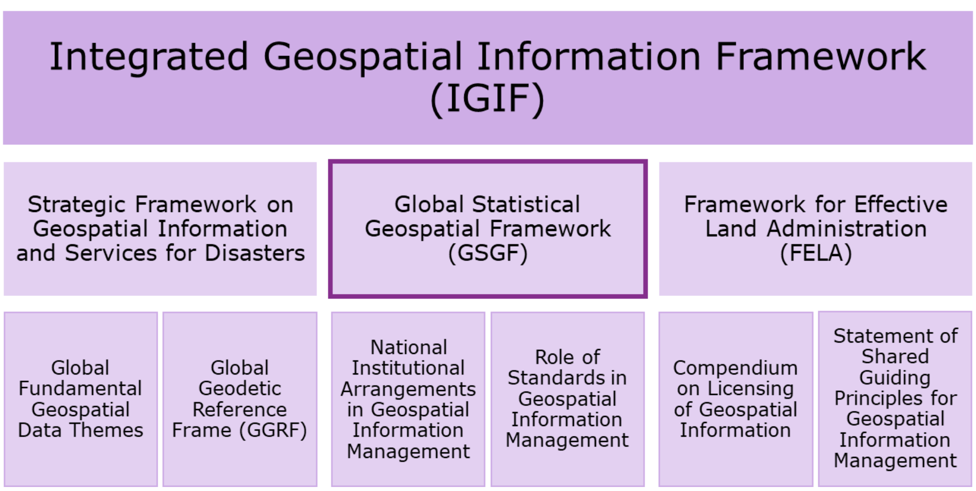

Figure 1. The GSGF under the umbrella of the Integrated Geospatial Information Framework (UN-GGIM, 2020a).

Third, the GSGF is part of the family of frameworks that aims to integrate data and information with the measuring, monitoring and reporting processes. For example, the Integrated Geospatial Information Framework (IGIF) encompasses other global frameworks, providing an umbrella for them, including the GSGF (see Figure 1). The IGIF and eight other related geospatial and statistical frameworks actively applied in Europe are presented in the third part of this document.

3. The Global Statistical Geospatial Framework (the GSGF)

The Global Statistical Geospatial Framework (the GSGF) is a high-level framework developed by the joint UNSC[1]/UN-GGIM[2] Expert Group on Integration of Statistical and Geospatial Information (UNSC and UN-GGIM, 2019), based on a national framework in Australia. Through the adoption by the United Nations Committee of Experts on Global Geospatial Information Management (UN-GGIM) in 2018 and the United Nations Statistical Commission (UNSC) in 2019, the GSGF has been recognised as a framework for the world that provides an underlying mechanism for integrating statistical and geospatial information.

The GSGF provides the international statistical and geospatial communities with a common framework to connect socio-economic and environmental data to appropriate locations and improves the accessibility and usability of geospatial statistics. The GSGF can thus be seen as a bridge between statistical and geospatial communities (Figure 2).

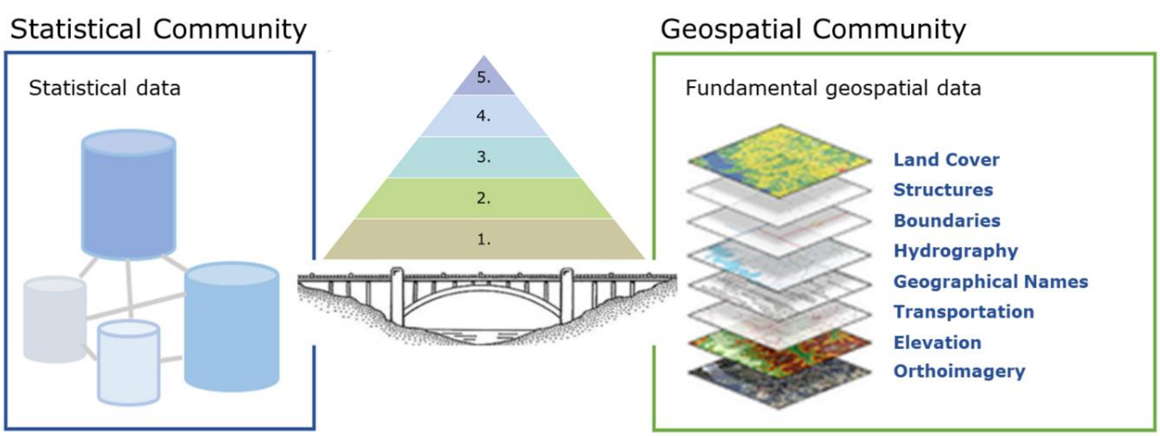

Figure 2. The GSGF aims to act as a bridge between the statistical and geospatial communities (derived from Brady, M., 2015).

[1] UNSC: https://unstats.un.org/unsd/statcom

[2] UN-GGIM: https://ggim.un.org/

The GSGF, is not in itself binding. Nor does it provide detailed implementation instructions. Like most high-level frameworks, it allows considerable flexibility concerning the "how". However, there is growing interest in and demand for implementation guidance for the GSGF, and in addition to GEOSTAT, the UN Expert Group (EG-ISGI) is working to provide implementation guidance at the global level (UN-GGIM, 2021).

Through the application of its five principles and four key elements supporting the implementation of the principles, the GSGF facilitates the production of harmonised and standardised geospatially enabled statistical data (Figure 3). Below, the five GSGF principles and their relations and interlinkages to each other are briefly presented. In Chapter 5, the implementation of the principles, including the four key elements, is studied in the European context.

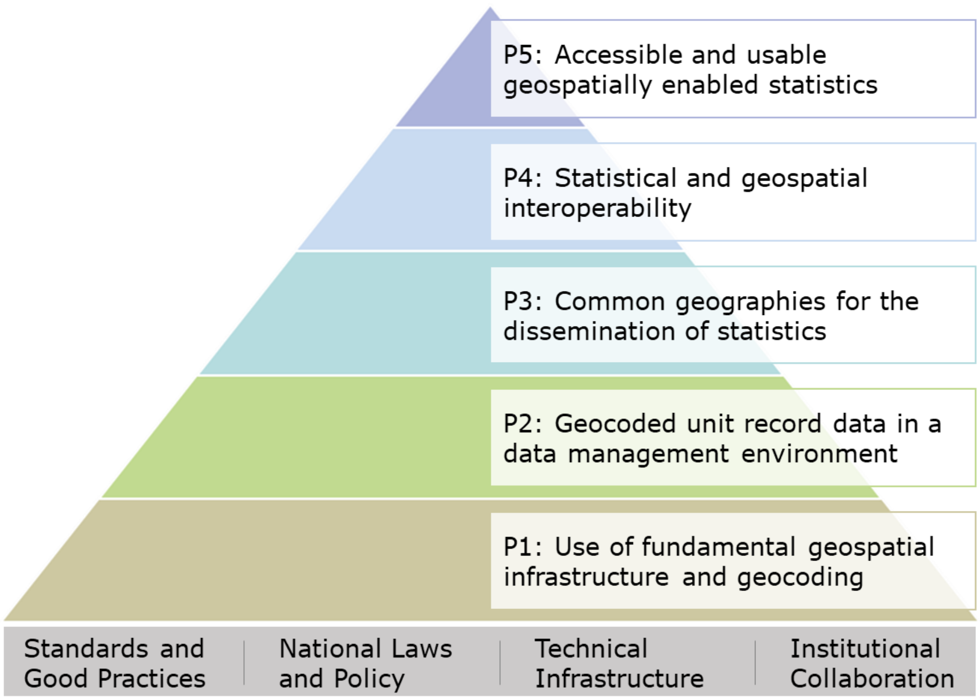

Figure 3. The GSGF consists of five principles and four key elements.

3.1 Develop Organised Geospatial Infrastructure

Principle 1, Use of fundamental geospatial infrastructure and geocoding, focuses on developing an organised national geospatial infrastructure that supports the creation of high-quality standardised location references such as a physical address, property or building identifiers, or other location description, and ensures the accurate assignment of coordinates and standard grid references. Preferably, location is recorded through the direct or indirect capture of x- and y-coordinates, and if available, a z-coordinate. This also includes the temporal components, time and date. Principle 1 is the basis for geospatially enabled information.

Table 1. The focus and objectives of Principle 1: Use of fundamental geospatial infrastructure and geocoding.

Focus	Objectives
• Information on location • A common and consistent approach to place each statistical unit in time and space using fundamental geospatial infrastructure	• Accurate, precise and consistent location information including address, property and building, meeting country-level agreed standards and good practices. • Geocoding results are as accurate and consistent as possible using common approaches or systems. • Any geocoding issues are consistently managed through application of standardised approaches.

3.2 Enable Geospatial Consistency in Data

Principle 2, Geocoded unit record data in a data management environment, supports the linking of each statistical unit record to a geographic reference, e.g. geographical coordinates or small geographical area, which will allow statistics to be applied to any geographical context. This is often referred to as geospatially enabling data, and it must occur within a secure, standards-based data management environment. This process applies the fundamental geospatial infrastructure from Principle 1. The implementation of Principle 2 supports the integration or linkage of data from other data sources, mitigating the challenges that arise with new geographies or changes in existing geographies.

Principle 2 includes the use of data management tools, techniques, standards and good practices to facilitate the linking and management of geocodes within statistical datasets. This also serves to ensure that privacy and confidentiality requirements are correctly managed for the released data.

Table 2. The focus and objectives of Principle 2: Geocoded unit record data in a data management environment.

Focus	Objectives
• Process to place and store • Data management environment • Privacy and confidentiality	• All statistical microdata are geospatially enabled. • Simplified data aggregation for common geographies • Adaptation to changes in geographies enabled • Effective data management • Clear data maintenance and custodianship roles • Geocoded information and metadata are consistent, interpretable and systematically maintained.

3.3 Ensure Common Geographies

Principle 3, Common geographies for dissemination of statistics, applies geography as a tool for integrating data. It uses a common and agreed set of geographies for the display, storage, reporting and analysis of social, economic and environmental comparisons across statistical datasets from different sources. Principle 3 also establishes the fundamental importance of balancing statistical and administrative geographies with other geographic referencing systems such as grids as a basis for establishing common geographies across datasets.

A common set of geographies ensures the consistent geospatial aggregation, comparability and dissemination of statistical data, irrespective of whether they are within gridded or administrative boundaries. Data are uniformly allocated to smaller administrative segments or statistical units such as mesh blocks that are divided according to political, property, or topological subdivisions, or consistently assigned to differently sized grid units, i.e. squares or pixels.

Table 3. The focus and objectives of Principle 3: Common geographies for the dissemination of statistics.

Focus	Objectives
• Geographies • Stakeholders' participation • Data aggregation	• Common geographies support data integration from different sources • Metadata support data aggregation, integration and use • Simplified visualisation, analysis and interpretation of information • Standard mechanisms to support conversion between geographies • Aggregation and disaggregation methods enhance quality, assessment, consistency, comparability and use of data

3.4 Enable Interoperable Data

Principle 4, Statistical and geospatial interoperability, urges the use of internationally adopted standards and good practices from the communities involved and identified in the Introduction. This enables greater interoperability of data, standards, processes and organisations. As a result, the potential application of a larger range of data and technologies increases, and thereby a wider range of information will be available and accessible for use in decision making. This will also lead to improved efficiency and simplification in the creation, discovery, integration and use of geospatial statistics and geospatial data.

Table 4. The focus and objectives of Principle 4: Statistical and geospatial interoperability.

Focus	Objectives
• Interoperability • Laws and policies supporting cooperation • Standardisation • Good practices	• Greater efficiency and simplification in the creation, discovery, integration and use of data • Service-based or machine-readable access mechanisms to provide greater efficiency of access and use, and to allow adaptation and evolution of uses over time • Application of a larger range of data and technologies

3.5 Provide Easy Access and Usability to Geospatially Enabled Data

Principle 5, Accessible and usable geospatially enabled statistics, highlights the need for data custodians to make geospatial statistics accessible and usable to ensure that data users can discover, access, integrate, analyse and visualise this information seamlessly for the geographies of interest. An important aspect of this principle is to ensure that data can be accessed using safe mechanisms, both to protect privacy and confidentiality and to support the analysis and evaluation of data in decision making. Other relevant issues include data quality and access to analysis, dissemination and visualisation capabilities.

The goal of Principle 5 is thus to ensure that geospatial statistics are accessible and usable for informed decision making. This is achieved by identifying, utilising and - where required - developing applicable policies, standards, good practices and technologies.

Table 5. The focus and objectives of Principle 5: Accessible and usable geospatially enabled statistics.

Focus	Objectives
• Data discovery • Data access • Dissemination • Visualisation and analysis • Data quality • Data maintenance • Metadata	• Data custodians can release data, considering privacy and confidentiality issues • Users can easily discover and access geospatially enabled statistics • Users can undertake analysis and visualisation • Enable machine-to-machine access and service-based integration of information • Users can know the status of modifications/changes in the data provided by data custodians

3.6 The Relations between the Principles

The five GSGF principles outline the broad processes by which a range of geospatial and statistical infrastructures and processes are applied to input data to enable integration. Here, Figure 4 presents the relations between the principles: Principle 1 is the foundational principle on which the other principles are built. It specifies the adoption of a common and consistent approach to creating high-quality standardised location references, building on a fundamental geospatial infrastructure. Moving to Principle 2, the geospatial infrastructure defined in Principle 1 is used to enable statistical data geospatially to the finest geospatial level possible.

The implementation of Principles 1 and 2 is a condition for the full implementation of Principle 3. This is to apply geography as a tool for integrating data, as the implementation of Principles 1 and 2 allows hierarchical and flexible aggregation into any output geography.

In turn, Principle 3 is an important condition for Principle 5, where common geographies form the basis for the dissemination of geospatial statistics. Common geographies should also be included in the National Spatial Data Infrastructure (NSDI) defined in Principle 1 for consistent use and output. Standardisation and interoperability issues in Principle 4 cut across all the other principles of the GSGF.

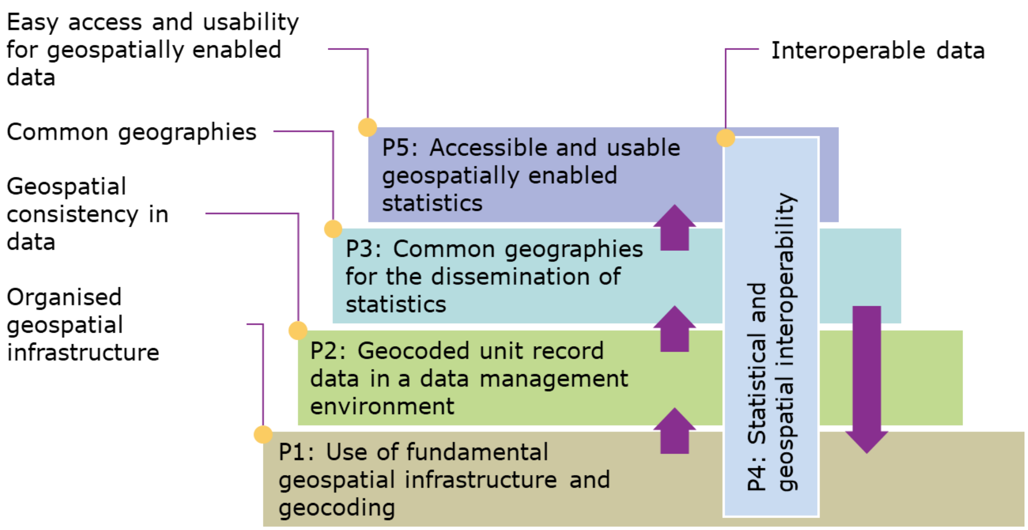

Figure 4. Relations between the principles.

4. The GSGF within the European Operating Environment

The five GSGF principles link to, and are implemented in, many levels within the European statistical and geospatial operating environment. In this study, we have examined the structure and operation of the statistical geospatial community, the interpretation of the GSGF's key elements and four different aspects that were identified as essential for further study in the European context. These four aspects are quality, data collection, confidentiality and innovation. Furthermore, several integration, standardisation and data-sharing frameworks are applied for the purposes of statistical and geospatial domains in Europe. The most relevant frameworks and their links to the GSGF are studied in Chapter 6. All these aspects together are used to describe and analyse the European operating environment in this document (Figure 5).

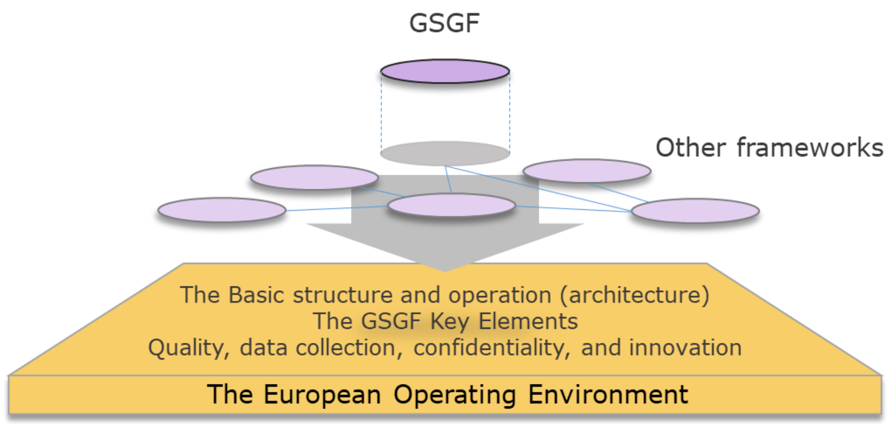

Figure 5. Linking the GSGF to the Operating Environment and other applied frameworks in Europe.

Describing the GSGF Europe Reference Architecture (GEOSTAT4, 2021) has been an important tool for the description and analysis of the European Operating Environment. This work is also based on the definition of requirements GEOSTAT identified as crucial in the implementation of the GSGF in Europe (Annex I). Each of these requirements is connected with a set of more specific recommendations that are addressed for particular actions. Some recommendations are also linked to a set of good practices from use case examples. However, in this document, the focus is on strategy and leadership activities. The requirements and recommendations presented in this chapter are based on this work unless otherwise indicated. Furthermore, material to support the implementation of the GSGF in Europe is provided on the EFGS website (EFGS, 2021).

In Europe, the organisations responsible for geospatial information and statistics work in elaborated legal frameworks, have strongly established working methods and use technologies that have evolved over many decades (Eurostat, 2019). The implementation of development frameworks therefore often faces different challenges compared to developing countries, for example. Well-established national data providers can serve their users according to their specific requirements and demands. This often means they each think and work separately in silos. As a result, data from different providers and countries are often difficult to combine and compare.

As the GSGF understands it, integration means real collaboration and knowledge exchange between the geospatial and statistical domains. The goal is that statistics, as well as all other data producers, use the same sources of geospatial information and the same geocoding services, resulting in a more efficient use of resources and improved integration of the data produced by different public organisations and the private sector.

On the other hand, strong institutions can benefit from transforming and adapting to new challenges (Eurostat, 2019). Some progress has already been made to ensure the smooth integration of different data. Within some countries, the combination of statistical and geospatial information is developing rapidly, based on close cooperation between their National Statistical Institutions (NSIs) and National Geospatial Information Agencies (NGIAs).

4.1 The GSGF Europe Reference Architecture - the Basic Structure and Operation

Effective collaboration starts by ensuring that the statistical and geospatial communities can share the same view of their operating environment and that they discuss the same concepts on common ground. Only then can solutions be built on the same common conceptual and theoretical base. To support this, the first steps have been taken in describing the GSGF Europe Reference Architecture (GEOSTAT, 2021a).

The first version of the reference architecture describes actors, roles, processes, services and concepts. Other reference architectures like CSPA, GSBPM, GeoGSBPM, GSIM and GAMSO have been considered and used in the building of the GSGF Europe Reference Architecture. The GeoGSBPM and GSIM frameworks are especially strongly related to this work. All the material about the reference architecture is available in the GSGF Europe Information Service on the web.

To reach the target state presented in this chapter takes time and resources if a system has not already been set up. Managing timely new and abolished reference data and microdata units also involves many processes and organisations, and requires agreements and updates in existing systems. It is recommended to start the work by defining the present status of the organisations and infrastructure and a clear common goal for the geospatial and statistical communities.

4.1.1 Data and Information

Data are at the heart of the GSGF. Several data streams feed the framework: geospatial reference data serving as the basis to geospatially enable statistics (Principle 1); statistical or other data to be geospatially enabled (Principle 2); and data on common geographies (Principle 3) to support the aggregation of geospatial statistics and the dissemination of the result (Figure 6).

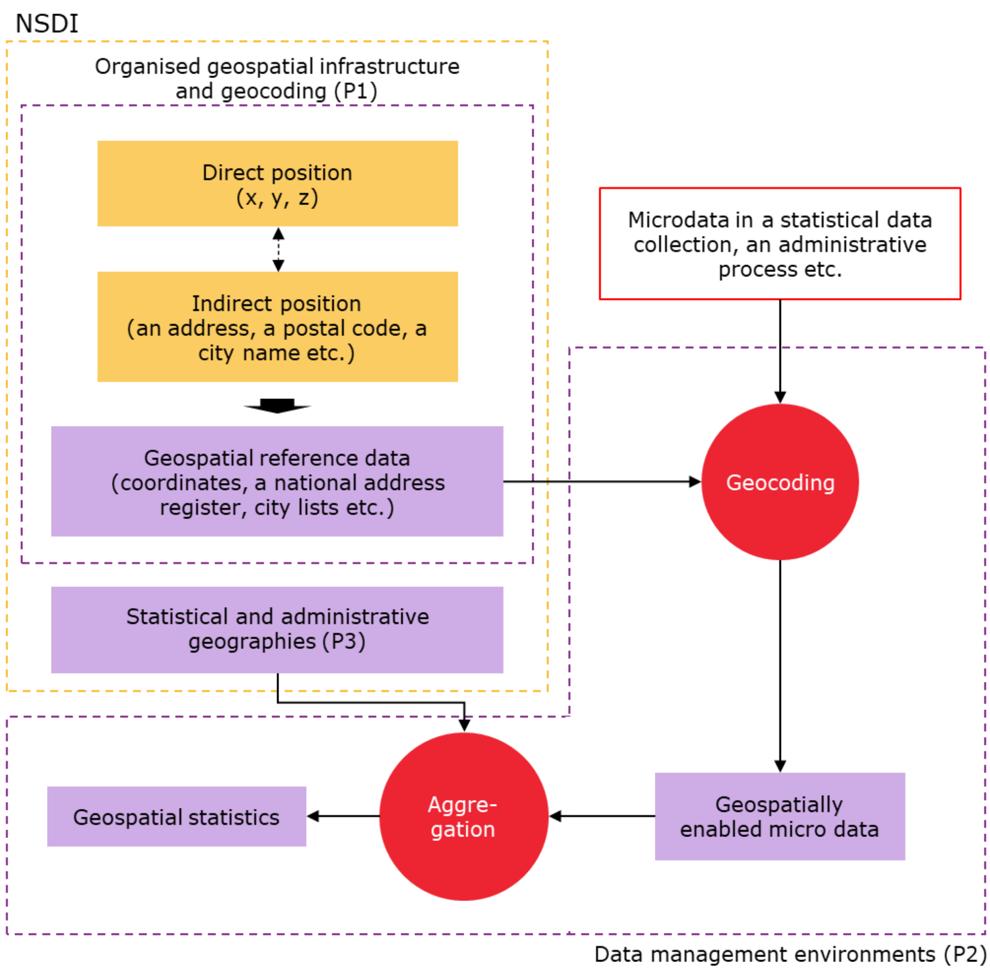

Figure 6. In this figure, the relationship between the geospatial infrastructure (P1) and the data management environment (P2) is illustrated. The NSDI is the common repository for geospatial reference data (e.g. address and building registers). The geospatial reference data serve as a geocoding infrastructure. The geocoding of statistical or administrative data results in geospatially enabled microdata. This in turn can be aggregated using common geographies (P3).

Principle 1 has a strong focus on geospatial reference data, where its content, structure and quality need to fit the purpose of statistical geospatial integration. Geospatial reference data, including services built on top of it, constitute the geospatial infrastructure. This infrastructure's critical role is to ensure a common and consistent approach to placing each data unit in time and space. Furthermore, GSGF Europe, as well as the global implementation guidance, urges the use of the most accurate geospatial reference data possible. This is reflected by the request to use point-based geospatial reference data for geocoding.

The request for a common geospatial infrastructure is not unique to the GSGF. Most European countries have established their National Spatial Data Infrastructures (NSDIs) in accordance with the INSPIRE directive. These NSDIs should comprise repositories for the geospatial reference data with a national standards-based address and building registers, including unique and persistent identifiers. NSDIs are currently built around authoritative geospatial data. There are good reasons for this requirement, as authoritative data provide a high degree of credibility and predictability. However, GSGF Europe suggests that NSDIs in the future will also need to include data from the broader geospatial community, including data from private producers and data derived from crowdsourcing and new technologies or sources, e.g. big data. These data sources are no substitute for authoritative geospatial data, but they can provide valuable enrichment and augmentation of the data ecosystem.

Principle 2 focuses on building an effective and secure data management environment. Success in the implementation of Principle 2 requires the active promotion of open access and use of authoritative geospatial reference data. For example, it is recommended that strong incentives for the whole of society be created to use the authoritative national address register, and that address data be released under open data licences.

Principle 3 aims to ensure the access and usability of contemporary and historical statistical and administrative geographies. The governance and management of common geographies include the agreed scale, reference dates, coding systems, resolution and accuracy of geographies. The framework of geographies also needs to be continuously developed in line with emerging user needs.

The interoperability Principle 4 highlights concerns how data travel from the source to the end user and is thus critical to the successful implementation of the GSGF. GSGF Europe provides a range of requirements and recommendations targeting data interoperability to enable the growth of data ecosystems. Semantic interoperability ensures that the precise format and meaning of exchanged data and information are preserved and understood: "what is available is understood". This includes syntactic aspects such as the terminology used to describe concepts, as well as describing the information's precise format. Both statistical and geospatial communities should clearly define a scope and resources for the application of common conceptual models, e.g. starting with the assessment of the GSGF Europe Reference Architecture.

The realisation of users' access to geospatially enabled data as requested by Principle 5 requires clear and simple data licensing policies, preferably releasing geospatial statistics as open data. Furthermore, European countries should intensify the development of dissemination solutions that enable users to customise the content and extent of data extracts. Systematic consultation with the geospatial statistics user community is required to understand and meet user needs.

4.1.2 Services

To develop smarter statistical geospatial integration and run efficient statistical geospatial processes, a set of modular services is necessary. These services would potentially be used by both statistical institutions and mapping agencies, but also by other organisations and external users. These services are linked to the statistical-geospatial process to perform different tasks.

In the GSGF Europe Reference Architecture, a list of generic services needed to produce geospatial statistics is identified and described. There are already widely agreed services such as services for the geocoding and dissemination of geospatial data, as well as services supporting the versioning of coordinate transformations and address services that are easily integrated with data collection platforms to increase interoperability.

An important step in the implementation of the GSGF is to provide access to open, interoperable and standardised geocoding services. These services should be open to the authorities and private sector so that they can rely on similar methods and tools for the geocoding and georeferencing inside and outside organisations to obtain consistent results.

4.1.3 Processes

The description of processes within the GSGF Europe Reference Architecture emphasises relationships between geospatial and statistical processes. Furthermore, the Geospatial Task Team under the Supporting Standards Group of the High-Level Group on the Modernisation of Official Statistics has developed a geospatial view of the GSBPM (GeoGSBPM) that emphasises those activities needed to produce geospatial statistics, using the framework of the GSBPM (UNECE, 2021). The results of this work are also integrated in the GSGF Europe Reference Architecture.

Currently, both statistical data collection processes and geospatial reference data management processes to produce geospatial statistics are typically carried out by statistical organisations quite independently. Geospatial reference data are often only used to geocode final statistical deliverables.

In the GSGF, the basic premise is that statistical processes depend on
the geospatial processes. These geospatial processes are the responsibility of NGIAs and are usually designed to produce geospatial reference data and tools for use by a wide community of data producers (Figure 7). All these statistical and geospatial processes, carried out by several organisations, must be designed to work together to ensure consistency in the production of location information. The target is defined by the GSGF and is expressed in Principle 1: "The process of obtaining location information and geocodes should use relevant, fundamental geospatial data from National Spatial Data Infrastructures or other nationally agreed sources." Furthermore, it is required that high quality, a standardised physical address, a property or building identifier or another location description are obtained at the microdata level. This ensures that accurate coordinates and/or a small geographical area or standard grid reference are assigned to each statistical unit.

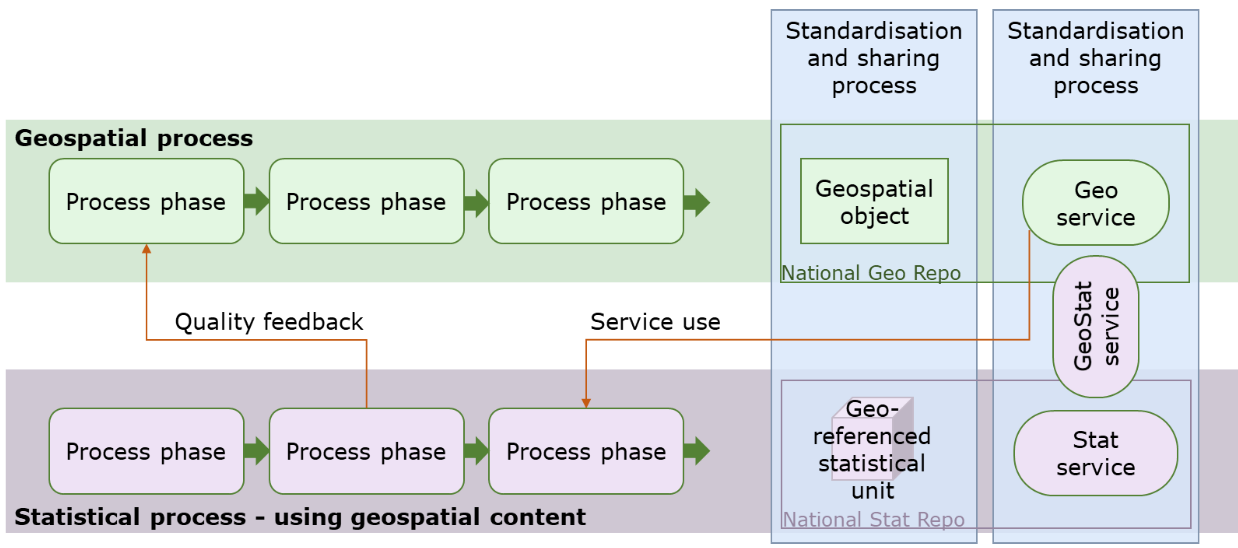

Figure 7. A general outline of a statistical process using geospatial content and geospatial processes, quality feedback and modular services.

An important observation is that the reciprocal dependence of the geospatial and statistical processes also requires the establishment of quality, as well as governance and management bridges between the two infrastructures to allow feedback between the process. For example, when the statistical process uses data and tools provided by the geospatial process, it is likely that the review and editing phase will lead to the detection of errors. In this case, it may be necessary to modify geospatial data based on the outputs of the statistical process using geospatial content to improve the quality.

Moreover, governance and management processes are needed to handle changes and workflows, and to support activities between the two infrastructures. The interactions between these processes can be complex. For example, some statistical geospatial processes may ultimately create new geospatial objects that should be added or changed in the geospatial reference data.

4.1.4 Actors and Roles

Each country seeking to implement the GSGF should first ensure that there is a clear agreement about the key stakeholders' roles and contributions especially NSIs and NGIAs. The description of actors and roles in the GSGF Europe Reference Architecture focuses on the identification, establishment and formalisation of generic roles that can be recognised across organisations at the national level (Figure 8). Actors can be linked to these roles and may vary from country to country, e.g. according to national and government bodies' administration and data custodianship assignments. The idea is that each country produces their own representation of the actors and their roles.

The use of data from NSDIs is required in Principle 1. This means that the roles and responsibilities of organisations and stakeholders involved in the production of geospatial information must be well defined, preferably through formal and binding protocols, agreements and Memorandums of Understanding (MOU). Furthermore, the implementation of Principle 2 requires clear data maintenance and custodianship roles to be defined. For example, custodian and stewardship models may need to be established to identify the most relevant stakeholders for a geospatial data source. Principle 3 extends this to coding systems and boundary data: clear custodianship roles for coding systems and boundary data for statistical and administrative geographies should also be defined at a national level to enable more efficient collaboration and ensure standardised data structures regarding these geographies. For example, coding systems and geographies are always produced, maintained and released from the same designated organisation and with the same data model.

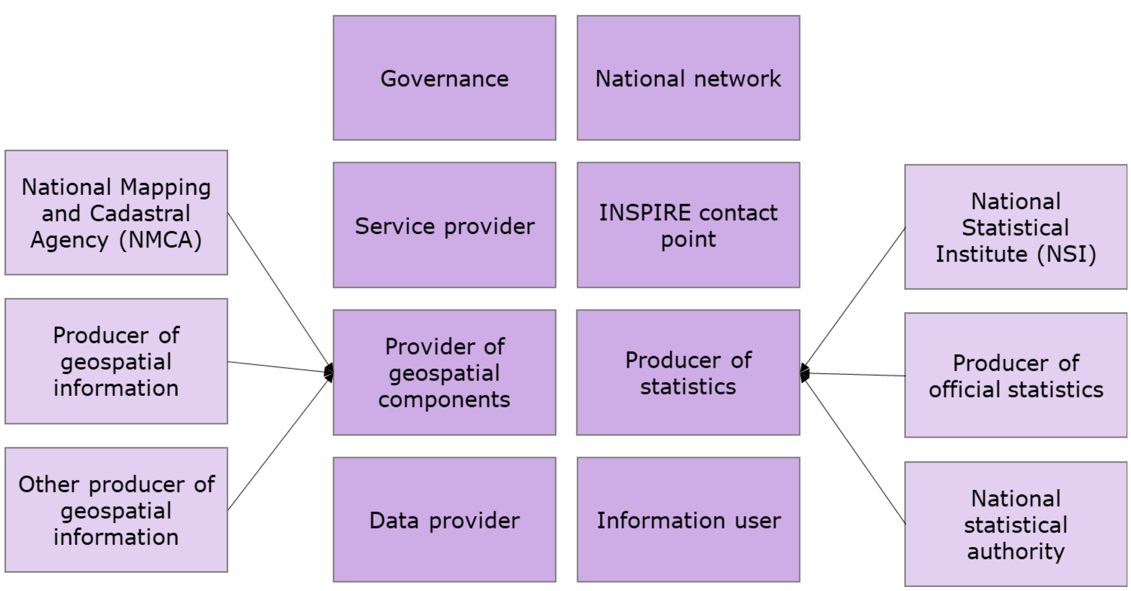

Figure 8. National key roles identified in the GSGF Europe Reference Architecture. An organisation may have one or more roles.

A number of initiatives and bodies of European collaboration contribute to the European statistical geospatial ecosystem with different roles (Table 6).

Table 6. European-level roles and actors identified in the GSGF Europe Reference Architecture.

Role	Actors
Provide statistics, geospatial information and other data on Europe	• Eurostat: Statistical office of the European Union • ESS Working Group on Regional, Urban and Rural Development Statistics • Eurostat and the ESS, through their cooperation and development programmes also carry out numerous activities and projects in non-EU countries to align their statistical system with the ESS.
Coordinate geospatial data production and services in the EU, production of pan-Union geospatial datasets	• GISCO: ESS working group on Integration of Statistical and Geospatial Information answers the needs of Eurostat and the European Commission for geographical information at the level of the EU, its Member States and regions. • Other ESS working groups, e.g. on quality • EuroGeographics
Designers of regional and territorial development /environmental policies	A set of national and international actors responsible for defining economic and social development policies at the level of regions or particular territories such as metropolises and urban or rural areas. European commission departments: • DG REGIO – Regional and Urban Policy • DG ENV – Environment • DG AGRI – Agriculture and Rural Development • DG CLIMA – Climate Action • Other policy DG's
Support of regional and territorial development /environmental policies	• EEA: The European Environment Agency • JRC: EU Commission Joint Research Centre
Fostering cooperation about global geospatial information management	Promote the development of global geospatial information, define a set of common principles among EU Member States, and between Member States and international organisations, and help European countries adopting these principles with geospatial information management. • UN GGIM: Europe – A subgroup of the United Nations initiative on Global Geospatial Information Management. Europe committee conducts actions to demonstrate the benefits of authoritative and trusted geospatial

	information with a focus on the achievement and monitoring of the Agenda 2030 SDGs.
European interest groups	• EFGS: European Forum for Geography and Statistics provides a unique voluntary network of experts for knowledge exchange and harmonisation of geospatial statistics. • NFGS: Nordic Forum for Geography and Statistics, NSIs and NGIAs, users, researchers and processors of geospatial statistics • EuroGeographics: A voluntary association of the Mapping, Cadastral and Land Registry Authorities of Europe for the development of the European Spatial Data Infrastructure.

The GSGF is a global framework and there are several other international-level roles that link to the contributors of international standardisation or cooperation, steering partners of regional development, designers of geospatial standards and interest groups. These include the following actors:

- The UNECE, which covers all the countries of Europe, works globally on modernising statistical methodology and statistical systems through the ModernStats programme, which also investigates the integration of statistics and geospatial information and closely cooperates with the ESS.
- The OECD: the Organisation for Economic Co-operation and Development has developed an ambitious programme to support cities and regions in developing, implementing and monitoring strategies to achieve the Sustainable Development Goals (SDGs).
- The United Nations Statistical Commission (UNSC), the United Nations initiative on Global Geospatial Information Management (UN-GGIM) and the related sub-groups (UN EG-ISGI, UN-GGIM: Europe) aim to play a leading role in setting the agenda for the development of global geospatial information and to promote its use in addressing key global challenges.
- The Open Geospatial Consortium (OGC) is an international voluntary consensus standards organisation, associating commercial, governmental, non-profit and research organisations worldwide so as to develop and implement of open standards for geospatial content and services, sensor web and Internet of Things, GIS data processing and data sharing.
- The Open Source Geospatial Foundation (OSGeo) is a non-profit, non-governmental organization whose mission is to support and promote the collaborative development of open source geospatial technologies and data.
- ISO/TC211: Standardization in the field of digital geographic information.
- The Global Forum for Geography and Statistics (GFGS) establishes a global network on geography and statistics, including NSIs, NGIAs and Research and Development Institutions

4.2 The Key Elements

When implementing a framework, such as the GSGF, not only technical issues need to be considered. Organisational, legal, semantic and other non-technical aspects require consideration, and a suitable strategy for handling them must be found. These aspects are to a large extent highlighted by the four cross-cutting key elements of the GSGF: Standards and Good Practices; National Laws and Policies; Technical Infrastructure; and Institutional Collaboration. These four key elements play an important enabling role that allows data to be obtained from the various sources and GSGF principles to be applied.

The four key elements intersect and intertwine, finally merging as the accessibility and usability of data in Principle 5. Thus, the produced geospatial data are readily available (accessibility), and it is ensured that these outputs are easily understood by a wide range of technical and non-technical users, and can be used to support policy and decision making (usability). This can only be achieved through active national and international collaboration.

4.2.1 Standards and Good Practices

The adoption of the GSGF requires that standards and good practices are applied within and across the geospatial and statistical communities. These include not only statistical and geospatial but also other standards and practices across the domains, e.g. information technology. In particular, Principle 4 urges the use of internationally adopted standards and good practices from both communities to enable greater interoperability, improved efficiency and simplification in the creation, discovery, integration and use of geospatial statistics and geospatial data.

The geospatial and statistical communities should be jointly involved in developing standards and good practices to consolidate and industrialise the implementation of statistical geospatial integration. For example, the use of data from National Spatial Data Infrastructures (NSDIs), discussed later in this chapter, is supported by the definition of technical standards and guidelines for geocoding in the NSDIs. The statistical community should also actively engage when new geospatial standards relevant for data integration are developed and be involved in the evaluation of utility for statistical production. Furthermore, the ESS, with Member States, should increase efforts to facilitate common tools, services, concepts and methods.

4.2.1.1 Standards

The effective leverage of the value of geospatial information requires the homogeneity of the delivered geospatial data and services. Given that geospatial information comes from many different sources and is managed by a large number of different providers, there is an overwhelming requirement to rely on a "common grammar" that ensures final deliverables sharing the same technical and logical components – that is to say, shared specifications. Standards play a key role in this respect and are essential to the delivery of authoritative geospatial services and products which meet the requirements of the wider community of users.

To ensure that this common grammar can be shared by the greatest number of actors and fully support public and private initiatives, it is important to favour open standards. Open standards are subject to full public assessment and use without constraints in a manner equally available to all parties. They are a central element in the growing open government trend. A goal of open standards is to foster interoperability – which is at the heart of Principle 4 of the GSGF. Standards are part of a national spatial data infrastructure, as they help to enable and improve the sharing, integrating and building of applications for handling geospatial information.

Many standards related to geospatial data and services are produced by international bodies like the Open Geospatial Consortium (OGC) and the Geographic information/Geomatics Technical Committee within the International Organization for Standardization (ISO). Geospatial standards can be classified in two categories: those related to data and metadata models (contents, concepts, topics); and those dedicated to technology and infrastructure (file formats, interfaces, APIs). Some examples of major standards are presented in Annex II.

The adoption of standards is a key element in assessing the level of maturity attained by an organisation in the production of geospatial statistics. For example, the "Guide to the Role of Standards in Geospatial Information Management", adopted by the UN-GGIM, groups standards according to "tiers of maturity" in geospatially enabled statistical processes. Each tier represents a series of steps in an organisation's ability to offer increasing levels of geospatial information and associated services (UN-GGIM, 2015).

4.2.1.2 Architecture

In the implementation of the GSGF, defining a reference architecture is a tool to support decision making and development work by defining the present and modelling the future status with the common goals of the statistical-geospatial domain. Architecture work helps remove silos, improves collaboration across organisations and ensures that technology is aligned to business needs. The architecture work thus helps describe how each organisation will meet future challenges and cooperate with each other. This work helps in standardising business processes, technology, applications and information in organisations.

Within organisations, this work can be supported by producing
reference architectures for selected domains. A reference architecture aims to provide a template conceptual and methodological solution that allows and supports the interoperability of implemented solutions and their sharing, as well as improved workflows within production processes. It also provides a common vocabulary for discussing implementations, often with the aim of achieving commonality. The international statistical community has some widely accepted reference architectures. The most used are

- CSPA
- GSBPM
- GeoGSBPM
- GSIM
- GAMSO

In respect of Principle 2, it is recognised as a requirement that a geospatial aspect be included in an organisation's business architecture. Location data should be recognised and fully integrated in the general data architecture of NSIs to facilitate the design and enable efficient workflows for data collection, integration and geocoding. It is also recommended to design and implement Geospatial Reference Architecture for statistics (Principle 4). This architecture should aim for the public administration of geospatial information and national interoperability.

4.2.1.3 National Spatial Data Infrastructures

National spatial data infrastructures (NSDIs) can be defined as frameworks consisting of geospatial data and services for finding, accessing, disseminating and using data, with related legislation, policies, institutional arrangements and human capabilities. Under INSPIRE, NSDIs have been put in place in EU Member States, and it is reasonable to establish a common geospatial reference data repository within these NSDIs.

The primary focus in the European context has been on technological solutions for discovering, sharing and using interoperable geospatial information. European "INSPIRE-like" NSDIs are often criticised for gradually falling behind current user requirements and representing too rigid public sector provider-centric frameworks, particularly outside the "silo" of geospatial professionals. One of the main challenges for the evolution of NSDIs over the next 5 to 10 years is incorporating and integrating new emerging spatial and non-geospatial data sources from public agencies, citizens and private industry with access that is as open as possible, following international standards. This will ensure that interoperability matters are considered. In addition, data needs to flow unharnessed across various domain-specific national and cross-border ecosystems. Semantic intelligence is also needed for converting information into knowledge. Future NSDIs should thus adopt a more user-oriented approach and design by maximising national geospatial data and capabilities, including dissemination mechanisms and the overall data access and usage. INSPIRE should thus enable the inclusion of more data beyond the geospatial data scope within the NSDIs by developing linked data flexible approaches towards a more statistical-geospatial data integration environment.

In practice, the five principles together aim to build a coherent NSDI and integrate statistics with it. The implementation of Principle 1 requires the production and use of common geospatial information, methods and tools for geocoding. In the European context, it is required that data from INSPIRE-compliant NSDIs are used. A well-established NSDI with a high-level of implementation and maturity builds a sound basis for an effective and secure data management environment (Principle 2).

The coordinated actions within NSDIs at national and organisational levels should trigger a need of common standards and interoperable data, services and technologies to support decisions and policy making at all the different scales for multiple purposes and application domains. NSDIs' main focus should rely on ensuring that geospatial data are more timely, flexible, publicly accessible and available at a minimum cost. At the same time, they ensure the enabling of the integration of relevant data from a wide variety of sources through location. Location matters, because it ties everything and everyone, assuming a 1 : n relation in which a single location reference is associated with several records, or objects about the built and natural environment.

4.2.2 Technical Infrastructure

Technical infrastructure, as the GSGF comprehends it, includes a broad range of technical capabilities spanning people skills, established and agreed methodologies and processes and systems infrastructure. The GSGF urges countries to consider how to embrace the framework within their statistical and geospatial production architecture by taking advantage of novel and emerging technologies and methods. This digital transformation is also accelerated by the EU through the enabling mechanisms of the 2030 Agenda. Innovative technologies such as the cloud, big geospatial data analytics, machine learning, geospatial knowledge services and integrated information systems can contribute substantially to this acceleration.

Improvements in technical capabilities are required. For example, where input data are derived from new technology and sources such as big data, more effective technical data and metadata management practices are required. Where service-based or machine-readable access mechanisms such as APIs are implemented. The data and tools should be open and free, wherever possible with no information loss due to technical or other interoperability issues.

In 2017, the European Interoperability Framework (EIF) was adopted (EC, 2017). EIF gives specific guidance on how to set up interoperable digital public services, and one of its four aspects is technical interoperability. This covers the linking systems and services of applications and infrastructures. The aspect also includes interface and service specifications, as well as data and metadata standards and formats.

One of the key messages of GSGF Europe is the importance of publishing data in open, interoperable and machine-readable formats and the use of services to automate the integration of data sources. Countries are urged to contribute to these efforts and to consider the GSGF as part of a set of wider efforts to modernise the production of statistics and geospatial information, as well as the general advancement of technology and information.

4.2.3 Laws and Policies

Laws and policies include key pieces of legislative, professional and social infrastructure that enable and in some cases constrain activities. They can include international and national data protection, privacy and confidentiality legislation, ethics and social licence requirements, open data policies and data access agreements. The GSGF primarily considers national laws and policies, but EU legislation must also be considered in the European context.

At European Union (EU) level, important initiatives are already asking for closer coordination between data-producing authorities and data users, and support the delivery of harmonised data (Eurostat, 2019). One of these is the EU action plan on e-Government and the subsequent Tallinn declaration on e-Government. The INSPIRE legislation requires data providers to share geospatial information about a wide range of Themes in a harmonised and standardised way, namely concerning data and metadata models. In addition, the European Statistical System (ESS) is working on a framework and methodology to promote and achieve harmonisation in data integration between countries via the GEOSTAT action. Specifically, the 2021 Population and Housing Census is a strong driver for geo-enabling statistics and integrating geospatial information with statistical production, as was ascertained in the first edition of the GEOSTAT action (GEOSTAT 1A, 2012).

It is important to be aware of the legislative, professional and social infrastructures that can both enable and constrain activities related to the use of geospatial information in the statistical and geospatial domains. Ethics is one of the topics with increasing interest. There are increasing expectations concerning the use of location, but at the same time, growing awareness about privacy and tracking location. For example, the UK Statistics Authority has published Ethical considerations in the use of geospatial data for research and statistics (UK Statistics Authority, 2021) with a geospatial ethics checklist. (Calder, A., 2021)

Legal restrictions and barriers should be reduced to free the users and producers of data as much as possible, and at the same time not to compromise data confidentiality and intellectual property issues. However, the first and imperative action is to find ways of securing a good understanding of the impact of an open and data-driven society. Only then can leaders commit and act to invest in the prerequisites of freely flowing integrated geospatial information; in terms of streamlined regulations, aligned policies and improved capacity building.

In Europe, the INSPIRE directive has provided the regulatory basis, whereas national policies have helped gradually open public sector geospatial data for free use in many countries. Data play an important role in the implementation and monitoring of European and global sustainable development programmes, UN Sustainable Development (UN, 2015) and the European Green Deal (EU, 2019). Statistical and geospatial data integration must support this transition of both public and private sectors and be prepared for the requirements of the future European data landscape and policies. The European Data Strategy (EU, 2021b), along with an initiative to create common sectoral European data spaces, is also being adopted. The goal is to create a single digital market. A key legislative instrument for driving the change is the new Open Data directive (revised Public Sector Information or PSI directive) which aims to secure extensive open access to public and private sector data so that companies and individuals who generate the data retain control. The directive will be implemented by a new act specifying High-Value Datasets (HVDs) that Member States will need to make available free of charge, in machine-readable format via APIs and/or bulk download. Geospatial and statistical data reside at the very core of HVDs. How the directives are and will be formally implemented in legislation and policies differs from country to country.

The benefits of data integration have been widely recognised in recent years by multiple stakeholders, including statistical and geospatial data providers. Despite this, European data production processes have been slow to change and respond, often within the constraints of complex institutional and legal contexts. Regarding the legal and political context, the absence of legislation and policies for cooperation in data integration, and the lack of economic funding and political incentives is also an issue that needs to be addressed (Eurostat, 2019).

However, several countries have established national strategies for the use of geospatial data. The strategies focus on the further development of NSDIs and show the direction of the work ahead. Important topics include national databases of geospatial information that meet a broad range of needs and framework conditions that are predictable and suited to the challenges in a digital society.

Good data quality is also essential for success in the implementation of Principle 2. There is a need to develop data quality assessment methods and approaches within the statistical geospatial production process, including all processes and sub-processes, particularly focusing on the geospatial domain, because it requires further and deeper work. Furthermore, good data quality could be enforced by developing a legal framework to support the validation of data at the source.

To foster the implementation of Principle 5, the European countries should support the implementation of clear and simple data-licensing policies. This should include the promotion of open data licences to release at least a core set of statistical variables for mid-resolution grids and the release of geospatial statistics as open data.

4.2.4 Institutional Collaboration

Europe is confronted by a mixture of societal, economic and environmental challenges such as climate change, ageing societies and massive immigration from developing countries. Understanding these major cross-border challenges and taking the right decisions requires new insights that we can only derive from new types of data and their combination. Unfortunately, organisational restrictions and a lack of cooperation between stakeholders need to be overcome first. The main issue related to institutional collaboration is the low level of interaction between geospatial and statistical European organisations, where there is either no or moderate cooperation between the data-producing authorities and a lack of coordination, common understanding and binding agreements (Eurostat, 2019).

The implementation of the GSGF requires capability building and the strengthening of institutional collaboration between the geospatial and statistical communities. The implementation is very much based on national cooperation and inspiration and the motivation of the two communities towards a common goal. Institutional collaboration can also be supported by formal agreements and/or positive institutional collaborative relationships, and should include outreach and education initiatives. After all, collaboration is beneficial for all parties.

In the European context, similar to several other regions, the strengthening of institutions and governance, as well as improved communication with stakeholders at all organisational levels, is critical to the implementation of the concepts around the integration of statistics and geospatial information presented in the GSGF. Closer cooperation between NSIs and NGIAs and other organisations with similar responsibilities, is needed. This requires commitment and the recognition of opportunities for cooperation, especially at the higher management and strategic level. It must also be recognised, both in the geospatial and statistical domains, what benefits the integration generates both for data producers and for users of statistics.

Nationally across communities and internationally across borders,
cooperation, exchange of knowledge and experiences and capability building, as part of concrete actions and projects, are important arenas for sharing knowledge about how to utilise and integrate geospatial information in statistics. Participating in and contributing to national and international events of both a formal and informal nature with representatives from both statistical and geospatial communities, will help connect people with valuable skills and knowledge to embrace the concept of integrating statistics and geospatial information.

In the GSGF, Principle 1 emphasises the building of formal working relationships in institutional agreements. This requires formal agreements between NSIs and NGIAs covering terms of access, licensing, governance and the use of geospatial information. Agreements may also need to involve other stakeholders such as municipalities or other sub-national bodies responsible for data provision. Data from the NSDIs need to be easily accessible and usable by NSIs or other public institutions at a low or affordable cost, but preferably free of charge.

Collaboration between statistical and geospatial communities at a European level is of vital importance. Institutions and initiatives should facilitate and support community-bridging forums, activities and projects such as UN-GGIM: Europe, EFGS[3] and Eurostat[4].

Regarding Principle 3 in the European context, it is recognised that there is a need to establish and maintain a consistent framework for national statistical and administrative geographies. For this work, geospatial agencies and NSIs should define clear custodianship roles for coding systems and boundary data at a national level, thereby enabling more efficient collaboration. They should also work together to improve the accessibility, speed of delivery and usability of national data on administrative and statistical geographies, with the aim of providing high-quality data under Open Data licences and obtained through OGC/INSPIRE-compliant Open APIs. Furthermore, geospatial agencies and NSIs should agree an approach on the scale, reference dates and accuracy of administrative and statistical geographies. This is relevant for ensuring data aggregation, data comparability, geographical coherence and spatial and temporal currency (up-to-date statistical data based on geographies' change).

[3] EFGS: https://www.efgs.info/

[4] Eurostat: https://ec.europa.eu/eurostat

INSPIRE has built a sound basis here too. A grid system is already settled within the INSPIRE principles (Statistical Units specifications), and it is a well-established feature among users of pan-European geospatial statistics. However, there remains a need to consolidate the use of existing statistical grid systems, introduce additional grid sizes, and explore the potential of the Discrete Global Grid System (DGGS) to achieve the next level of data comparability. Furthermore, the potential of common statistical and administrative geographies, harmonised methods and guidance should be explored. This should all be done jointly by the statistical and geospatial communities.

Future NSDIs should also challenge institutions. They should directly address implementation requirements. More effective implementation can be addressed by formulating enhancing governance, encouraging multilevel participation, and being open to new organisational structures and models. This will only be possible with changes in the institutional context, including processes of institutional building and the definition of roles and responsibilities.

The parallel use of national, EU and global grid systems for statistics should be coordinated in the ESS to ensure that all grids are developed with comparable quality and disclosure control. Furthermore, all NSIs within the ESS are also encouraged to follow and support the development of EU legislation to support the development of the European framework of statistical geographies, in particular by means of the implementation and further development of the territorial classifications defined in the TERCET Regulation (Eurostat, 2018).

Principle 4 urges statistical and geospatial agencies to ensure semantic interoperability across data domains, for example, ontologies for addresses and buildings by working together to inform data providers on the need for the conceptual harmonisation and benefits of data interoperability. Both statistical and geospatial communities should also clearly define a scope for the application of common conceptual models, engage in work of relevant national and international working groups and assess the existing conceptual models. They should also ensure that there is clear assessment, planning and common agreement on priority areas for action and an agreement on roles and contributions by key stakeholders, particularly NSIs and NGIAs. Furthermore, geospatial workflows are improved within statistical production by developing and sharing common and reusable tools and data integration service solutions to be shared in open repositories. Principle 5 also looks after collaboration and an EU official geospatial statistics portfolio, for example, could be developed there based on a user needs analysis.

Access to open, interoperable and standardised geocoding services, preferably APIs, is emphasised. National geocoding services should be open to the authorities of other Member States to rely on similar methods and tools for the geocoding inside and outside countries and to obtain consistent results. For example, this would allow the recording of workplace addresses of citizens working abroad producing more spatially precise and reliable statistics. Common services could provide a better basis for cross-border geocoding, hence improving the calculation of statistics on cross-border commuting and migration, including labour and accessibility statistics, e.g. studies of Labour Market Areas (Eurostat, 2020) and A Walk to the Park (European Commission, DG REGIO, 2018).

The European bodies, national geospatial agencies and NSIs should work closely together on developing and publishing scalable tools and generic geospatial services to standardise production processes. They should also underpin governance and guidance on the provision of these services. This will enable more flexible and open statistical geospatial integration, providing more value and interactivity for users.

Thus, any geospatial information should be built on agreed, authoritative and INSPIRE - compliant location reference data for geocoding and services. To enable this, the roles and responsibilities of organisations involved in the production and maintenance of geospatial information need to be well defined through formal protocols, agreements and Memorandums of Understanding. Custodian and stewardship models may also be needed to identify the most relevant stakeholders for a geospatial data source. A common and consistent approach for both public and private actors will thus be developed.

Institutional collaboration is supported by several organisations and interest groups to foster and support cooperation at both national and international levels. They provide a platform for breeding ideas and spreading knowledge and practices. One of them, the European Forum for Geography and Statistics (EFGS), provides a unique voluntary network of experts for knowledge exchange about and harmonisation of geospatial statistics.

4.3 Going beyond – Topics for Future Enhancement

In addition to the aspects studied above, four topics for the future enhancement and extension of the original scope of the GSGF were recognised. These are quality, data collection, confidentiality and innovation.

4.3.1 Quality

Within the European Statistical System (ESS) a certain quality framework was developed in order to ensure that the European Statistics produced are in accordance with commonly accepted standards by using at least good practices. One of the main objectives is to enhance trust in official statistics as a solid decision evidence-based fundament for users and policy makers. The elements of the framework are provided by:

- The European Statistics Code of Practice (CoP) sets the standard for developing, producing and disseminating European statistics, along the lines of the institutional environment, statistical processes and statistical outputs. Since its last version (2017), the CoP includes the Quality Declaration of the ESS, establishing the commitment of the ESS as a whole to the production and dissemination of independent high-quality statistical information at European, national and regional levels.
- The ESS Quality Assurance Framework (QAF) complements and breaks the CoP further down, identifying possible methods, tools and good practices that can provide guidance and evidence for the implementation of the CoP, at process and institutional levels.
- A multi-dimensional quality definition anchored in article 12 of the regulation Quality in European Statistics defining the quality of official statistics by the following criteria
  - 'relevance', which refers to the degree to which statistics meet current and potential needs of the users;
  - 'accuracy', which refers to the closeness of estimates to the unknown true values;
  - 'timeliness', which refers to the period between the availability of the information and the event or phenomenon it describes;
  - 'punctuality', which refers to the delay between the date of the release of the data and the target date (the date by which the data should have been delivered);
  - 'accessibility' and 'clarity', which refer to the conditions and modalities by which users can obtain, use and interpret data;
  - 'comparability', which refers to the measurement of the impact of differences in applied statistical concepts, measurement tools and procedures where statistics are compared between geographical areas, sectoral domains or over time;
  - 'coherence', which refers to the adequacy of the data to be reliably combined in different ways and for various uses.

The implementation of the ESS Quality Framework is further underpinned by statistical regulations and facilitated by quality-related standards, guidelines, methods and tools, providing particular focus on quality assessment and reporting activities.

The production of geospatial output is in some way specific as there is a focus on certain processes and other input sources. Furthermore, there might be different very specific user groups and user needs for geospatial products. On the other hand, the habits and interests of users of statistical figures have changed. Caused by the technological progress users are able to process not only tables but other appearances of statistical outputs as well. Because of this, nowadays nearly every statistical product is accompanied by geospatial outputs.

Given the fact that the demand for and importance of geospatial
outputs is continuously increasing, the question is how much the elements of the European quality framework are affected when we deal with geospatial processing and outputs.

The Global Statistical Geospatial Framework (GSGF) facilitates the integration of statistical and geospatial information. A Framework for the world, the GSGF, enables a range of data to be integrated from both statistical and geospatial communities and, through the application of its five principles and supporting key elements, permits the production of harmonised and standardised geospatial products. As the major framework the GSGF is targeting three major communities, or groups of stakeholders, and their different roles in terms of data provision and data integration: the geospatial community, the statistical and administrative data communities. The GSGF also stresses the importance of the integration of new sources of data for the production of high-quality geospatial information.

The GSGF contains in total 18 requirements classified related to the five principles. When looking at the requirements they can be interpreted as a kind of self-standing quality framework. The question however is how far the requirements encompass, as well the quality criteria of the ESS.

A mapping (Table 7) between the requirements of GSGF Europe (Annex I) and the ESS quality framework (quality criteria and Code of Practice principle) shows that there is no direct 1-to-1 relation between GSGF principles and the components of the ESS framework.

In Principle 1, accuracy is addressed by requiring high quality, standardised physical address, property or building identifier, or other location description, in order to assign accurate coordinates and/or a small geographic area or standard grid reference to each statistical unit, i.e. at the microdata level. Furthermore, comparability and coherence can be assured by using a national spatial data infrastructure. It is also interesting to note that the very recently introduced principle 1bis of the CoP (cooperation and coordination) is very much supported by GSGF Principle 1 when looking at the requirement "Build formal working relationships on institutional agreements".

Principle 2 of the GSGF is the most relevant driver for accuracy since the geocoding process is the main focus. Very detailed aspects of the data and how accurate they should be are found. Since linkage of data is addressed as a key process there, linkage errors as typical aspects of non-sampling errors are the predominant aspects here. By requiring a well-defined approach for non-matches, a commitment to reduce processing errors is clearly visible.

Table 7. Mapping of GSGF requirements against the quality framework of the ESS.

Quality criterion / GSGF Europe Requirement	Relevance	Accuracy	Timeliness	Punctuality	Accessibility and clarity	Comparability	Coherence	CoP Principles
Principle 1. Use of fundamental geospatial infrastructure and geocoding								
Use data from National Spatial Data Infrastructures						x	x	1-8, 12,14
Use point-based reference data for geocoding		x						7,8,12
Build formal working relationships on institutional agreements			x	x				1-3
Principle 2. Geocoded unit record data in a data management environment								
Build an effective and secure data management environment		x						5
Include a geospatial aspect in organisation's enterprise architecture		x	x		x			2, 13,15
Ensure consistency and quality of geocoding results		x						7,8,12
Use point-of-entry validation in collection of administrative or statistical data		x						3,8
Define common data quality frameworks taking into account spatial and temporal consistency	x	x	x	x	x	x	x	4,11, 12,13, 14,15
Principle3. Common geographies for production and dissemination of statistic								
Set up and maintain a consistent framework of national statistical and administrative geographies			x			x	x	13,14
Consolidate use of existing statistical grid systems and explore the potential of evolving global grid systems	x	x				x	x	13,14
Principle 4. Statistical and geospatial interoperability – Data, Standards and Processes								

Improve geospatial workflows within statistical production								7,8
Enable data integration through consistent semantics and concepts across domains							x	14
Publish data once and leave them at their source to be reused many times	x				x			15
Increase use of services and semantic web technology to enable innovation in a wider data ecosystem	x				x			-
Principle 5. Accessible and usable geospatially enabled statistics								
Implement clear and simple data licensing policies					x			6
Use service-oriented data portals supporting dynamic integration of data	x				x			7,15
Define clear national and European rules to ensure protection of privacy		x			x			5
Facilitate data search and use through cataloguing and improved guidance	x				x			15

The implementation of an effective and secure data management environment also serves the purpose of increasing and ensuring punctuality and timeliness.

Taking advantage of common geographies for production and dissemination of statistics as stated in Principle 3 is a useful approach to bring users in a situation with comparable results in geospatial outputs. The request that a geometry of boundaries, data on national statistical and administrative geographies should be available to all users no later than six months after the reference date when changes occur is a clear statement to accept a standard regarding timeliness.

Principle 4 can be seen as the key principle for staying relevant. It recognises that both communities (statistical and geospatial) operate their own models and standards but also emphasises the need to overcome structural, semantical and syntactic barriers in order to improve the discovery, access and use of geospatially enabled statistical data. Especially communities using open data are formulating users' needs which are relevant for the production of geospatial outputs.

Accessibility and clarity are the main topics of Principle 5, where the implementation of platforms and services to which users can return when interested in geospatial outputs are addressed. Service orientation standardisation of tools and continuous improvement of usability are key elements for statistical services in general and in Principle 5 they are addressed properly.

To summarise, it can be said that the GSGF, while not directly addressing the ESS quality dimensions, generates a high degree of interoperability between the two concepts. Although there might be overlaps in a sense that many of the GSGF requirements can be assigned in a way to more than one quality criterion but implicitly it can be concluded that by adhering to the GSGF requirements a quality assurance process relevant for the quality in the ESS is implicit.

As a consequence, elements in the GSGF could be taken as a possible donor when it comes to providing contents for quality reports, for instance, based on the European quality reporting structure SIMS (Single Integrated Metadata Structure). Work package 3 (Quality) of GEOSTAT 4 made a first attempt to suggest possible quality elements for geospatial outputs and provides a list of possible quality indicators which can also be related to GSGF principles.

4.3.2 Data Collection

Statistical processes are based on inputs from diverse data sources, from which the produced indicators and dissemination files are the outputs. Traditionally, statistical operations have been built on survey data, named a sample. They also rely on administrative data, like registers, or social or fiscal declarations. In recent years, access to alternative sources has become widespread, with the possibility of enhancing information held by private actors on a more regular and detailed basis - for example, banking information. Big data from the digital economy are also a focus of the statistical community in modernising official statistics production in the social, economic and environmental development dimensions.

However, survey data and big data correspond to very different kinds of collection processes. They present common challenges from a geospatial perspective. This aspect is illustrated by two common challenges: first, the quality of the geocoding process; and second, the representativeness of the data at a local scale. The former is more related to the new data paradigm encompassing innovation and requiring more exploratory approaches. The latter is more traditional, with a well-established knowledge body and long tradition within the statistical production process.

The first challenge concerns the geocoding process that used to produce geospatial statistics from unit data and its quality assessment. GSGF Europe recommends the use of a common base of location data for statistical and administrative data through NSDIs. It highlights the importance of consistent and standardised methods and tools for locating information and ensuring the assignments of accurate location information. Hence, NSIs and NGIAs must establish and disseminate technical standards and guidelines for geocoding to develop a common approach for all producers and users. Such guidelines may state which type of location data should be used for which cases, depending on data type, the kind of rules to deploy if locations are missing, and how to improve matching between records and help ensure the result will be the same regardless of the person or institution conducting the geocoding. Technical standards and guidelines must meet quality criteria, e.g. quality framework, point-of-entry validation, etc. The quality assessment should address all geocoding processes from data collection, metadata and production to dissemination.

In this regard, survey data need to have a more comprehensive understanding of the geographical space. This is accomplished mainly through a point-based geocoding infrastructure for statistical production, a key condition supported by NSDIs. The inclusion of the geospatial dimension in the GSBPM collection phase, including survey operations, provides a more accurate location for the statistical units. The survey sampling design and production of geospatial statistics is thus improved on any small areas.

In the case of big data, the location is often collected natively, while the records are captured. This is expecially the case for smart censors' data. This geocoding is natively dependent on the technical parameterisation of smart sensors. This means that the technical and methodological choices occur before the statistical pipelines, which must deal with information as it is. Hence, there is a specific operational concern to promote common technical standards about geocoding from smart sensors – or the geospatial processes need to be adapted to the location references provided.

The second challenge is related to the representativeness of the available data at a local scale. Survey data are collected according to a sampling frame that is specifically established to ensure that the surveyed units represent the whole population to be covered, for a given geographical scope and scale. Here, the challenge is to collect enough answers at a very local scale. If the response rate is too low, the results obtained do not include enough records to be able to produce reliable geospatial statistics (too few geographical objects and a lack of spatial – and attribute – completeness). This problem can also arise in the event of partial answers or missing values (compromising dataset completeness). In this case, certain statistics on particular variables will only be available at a macro level (regional and national), without being produced in local areas. This will lead to inaccurate generalisation, affecting spatial analyses and resulting in extrapolated and misleading data representation and conclusions (ecological fallacy). This will ultimately lead to misinterpretations of the geospatial statistics.

Concerning big data, the large number of observations does not guarantee the representativeness of these recordings. There may be coverage bias, closely linked to social or economic characteristics for sub-groups of population. For example, mobility statistics produced from mobile phones do not cover people who do not have one. The coverage of collected data in relation to the population of interest must be clearly identified. Attention must be paid to measuring the representativeness at the local scales and its feasibility for producing geospatial statistics. For example, there is a risk of spatial bias if smart sensors used to collect data are not homogenously distributed in the territory.

For both survey and big data, GSGF Europe recommends the building of a measure of "non-coverage" that takes the target level of geography into account, including possible geospatial collection bias and the relative importance of missing values at a local scale to measure the level of (spatial and attribute) completeness, as well as correction methods. This measure of "non-coverage" should also take the possible effects of the dissemination and interpretation of geospatial statistics into account.

Survey data and big data are addressed in Principles 1, 2 and 4. When using survey data as opposed to big data, there is a greater scope for action, especially from the specify needs phase to the collection phase of the statistical production process. This is particularly visible when designing the point-based data infrastructure and in the geocoding process (Principles 1 and 2). This is why it is so important to define and disseminate guidelines on data models and metadata for private actors when providing (big) administrative data sources.

4.3.3 Confidentiality

Statistics with a geospatial component presents additional requirements in the management of statistical confidentiality. The risks of disclosure are increased by the information's geographical dimension through a process of "differencing" results. This section therefore sheds light on the management of data confidentiality, which is taken into account in the principles and recommendations of GSGF Europe.

Statistical data are subject to two data protection frameworks (Eurostat, 2021a, 2021b):

- The general personal data protection framework applies every time information about individual persons is collected, for whatever purpose. The General Data Protection Regulation (EU, 2016) strengthens the rights of data subjects and the obligations of controllers. The personal data protection aspects – data security, data traceability and data access – should be an essential element of the design of any data collection.
- The protection of data collected for statistical purposes – "statistical confidentiality" – is a fundamental principle of official statistics. Statistical confidentiality means that data on individual persons (or business entities) may be used only for statistical purposes, and that rules and measures will be applied to prevent the disclosure of information concerning an individual person or business entity.

Confidentiality issues concerning geospatial data are much the same
as confidentiality issues concerning any kind of data that are directly linked to individuals (microdata) or that present a risk of reverse identification of individuals (sensitive data). Yet the unique characteristics of georeferenced data present particular concerns and challenges.

The specificities of geospatial data have been studied by the UN Expert Group on the Integration of Statistical and Geospatial Information (UN EG-ISGI) with a set of guidelines. This document is aligned with these findings, which are also included in a proposal to extend the Quality Assurance Framework (ESS, 2019) with geospatial aspects.

First, geospatial data provide new options for revealing individual characteristics, including location and geographic footprints, which are precisely based on the spatial dimension of the phenomena under study or of the data collected:

- The geographical characteristics may give rise to special socio-economic relationships among individuals with strong spatial autocorrelation. This accords with Tobler's first law of geography (Tobler, W.R., 1970), which states that "Everything is related to everything else. But near things are more related than distant things." Thus, spatial autocorrelation refers to the pattern in which observations from nearby locations are more likely to have similar features than those from distant locations. For areas with a low number of observations, which are often those where the density is weak, the disclosure risk is therefore higher.
- The disclosure risk also increases for geospatial data because of geographic differencing issues, which occur when the same data are disseminated in different non-nested or similarly shaped geographies. In some cases, attributes can be deduced for a number of units below the threshold by subtracting the counting of an area from the counting of another enclosing area. Therefore, by using this method, data for the area can be obtained that is not common to both areas, which may result in a privacy breach.
- Trajectory geospatial data are doubly exposed, insofar as the geography is doubly combined (point of departure and point of arrival), and the risk increases with every additional location added to the journey.

Second, the statistician's toolbox does not yet include a suitable device for the management of the statistical confidentiality of geospatial data. Dealing with geospatial data adds a layer of complexity to the disclosure control process, because it requires the implementation of specific methods that need greater computing power than is available with classic computers and traditional methods.

Third, the growing demand for geospatial data dissemination
reinforces risks of unit identification. On the one hand, there is a general movement for open sourcing the fundamental information that constitutes part of the National Spatial Data Infrastructure like address registers, cadastral parcel boundaries, geolayers used for building informative maps, etc. Making these tools available is aligned with GSGF principles but may provide proficient tools to help identify statistical units. On the other hand, there is an increasing focus on spatial impacts of public and private actions, meaning that policymakers, analysts and even citizens are seeking detailed information across a broad range of spatial dimensions at a very detailed scale, which may again risk the disclosure of individual characteristics.

Major guidelines are under way by EG ISGI and they appear to be aligned with GSGF principles and recommendations can be summarised as follows:

- Increase the level of awareness of geospatial data's specifications for the management of confidentiality at the global, regional and national levels.
- Acknowledge these specifications in national statistical or privacy laws, data protection and release policies, nationally agreed guidelines, national, regional, or global quality assurance frameworks.
- Foster the collaboration with the scholarly community and other official bodies to promote research and explore new paths.
- Define a policy in which disclosure control methods must be applied to geospatial statistics in accordance with both the features of data and to dissemination targets.
- Include the monitoring of actual disclosure control in Quality Assurance Framework
- Foster the integration of the spatial dimension within existing software for the management of confidentiality.

Confidentiality is at the core of the successful implementation of Principle 5. An important aspect is the definition of clear national and European rules to ensure the protection of privacy. For example, to disseminate national grid data or small area statistics, countries should define, describe and publish their own principles for the preservation of privacy with respect to existing national legislation and policy.

4.3.4 Innovation

The evolution towards a broad portfolio of geospatial statistics is demanding. The need to build in-house capability in methods and tools will vary between organisations depending on ambitions and the needs specified for statistics. For good reasons, establishing competence and resources will often start in a small specialist environment and within an exploratory approach, but the concept of the integration of statistics and geospatial information should eventually be communicated throughout the organisation and with collaborating organisations. Within NSIs, this needs to occur through systematic collaboration between geospatial specialists and methodologists and statistical production units in the design of statistical production processes is of the utmost importance. Another element in the strategy may include organisational changes and new ways of managing statistical production. Ultimately, the increased use of geospatial information in the statistical production process will require geospatial specialists and production teams to be brought together.

Emerging technologies provide novel methods for the effective collecting, processing, managing and integration of exponentially growing quantities of multidisciplinary geospatial and other data. These developments need to be monitored and adopted by European NGIAs and NSIs to be able to provide up-to-date information and added value services to a wide range of users.

National and EU-level policies and legislation can create serious challenges for the effective use of new technologies – for example, in terms of strict privacy and security regulations and overly rigid and outdated specifications and processes for data acquisition, integration and management. In addition to legislative reform, the successful adoption of technological innovations requires holistic support through financial investment, policy intervention and capacity building to deliver adequate improvements for society as a whole.

However, in many cases, the innovation can be rather obvious and "low-hanging" rather than radical and disruptive. In the following, three themes for assisting the collaborative efforts of NGIAs and NSIs to promote the integration of geospatial and statistical data are discussed:

- Cross-cutting collaboration throughout society: Open-minded agile collaboration across and beyond government is needed to fully map and understand user needs, capacities and the potential to integrate geospatial and statistical information. This requires low-threshold coaction between public agencies, researchers, businesses and civil society – for example, by establishing a national forum for identifying high-value use cases, pinpointing common barriers and deficiencies and promoting improved ways of integrating geospatial and statistical data.

- Experimental culture and facilities for innovation: An innovation culture within organisations and across governmental domains can be created and nurtured by visionary and empowering leadership and embracing new ideas, providing space and time for experimentation and tolerating risks. To facilitate experimentation, decision makers need to establish innovation programmes and local/national hubs. They also need to provide access to R&D cloud platforms to enable shared co-located "playgrounds" for studying novel methods of data processing, analysis, visualisation and dissemination. Staff should be granted dedicated time allocations for exploring "freestyle" in collaboration with other information communities and users.
- Monitoring and aligning business processes – from intra-institutional silos to inter-institutional dataflows and services: NGIAs and NSIs usually plan, implement, monitor and improve their business processes considering only their internal organisational set-up and customer needs. However, a major part of (geospatial) dataflows, products and services could be streamlined and matched to ensure NGIAs and NSIs could harmonise and seamlessly share common processes. In practice, this would require the reviewing of data content, services and processes at different levels by using common enterprise architecture (EA) tools and methods. As a result, common data products, services and procedural interfaces could be identified and designed to ensure any duplication of work and services and Open APIs, as well as other inefficiencies, barriers and sources of unnecessary costs, could be uncovered and removed.

In GSGF principles, there is a need, linked to Principle 4, to increase the use of services and semantic web technology to enable innovation in a wider data ecosystem. Furthermore, Principle 5 encourages the use of various mechanisms to facilitate data searches and use. Here, user-centred product design might be a method to better meet the user requirements of geospatial statistics. The ESS should also develop efforts for an EU official geospatial statistics portfolio based on user needs analysis to better respond to challenges.

5. Surrounding Frameworks on which the GSGF Builds

There are several European integration, standardisation and data-sharing frameworks for the purposes of statistical and geospatial domains. As the GSGF bridges these two domains, there is a need to understand the role different frameworks play in statistical and/or geospatial fields, and how they are linked to the GSGF. Nine frameworks are studied here. They are presented in the figure below.

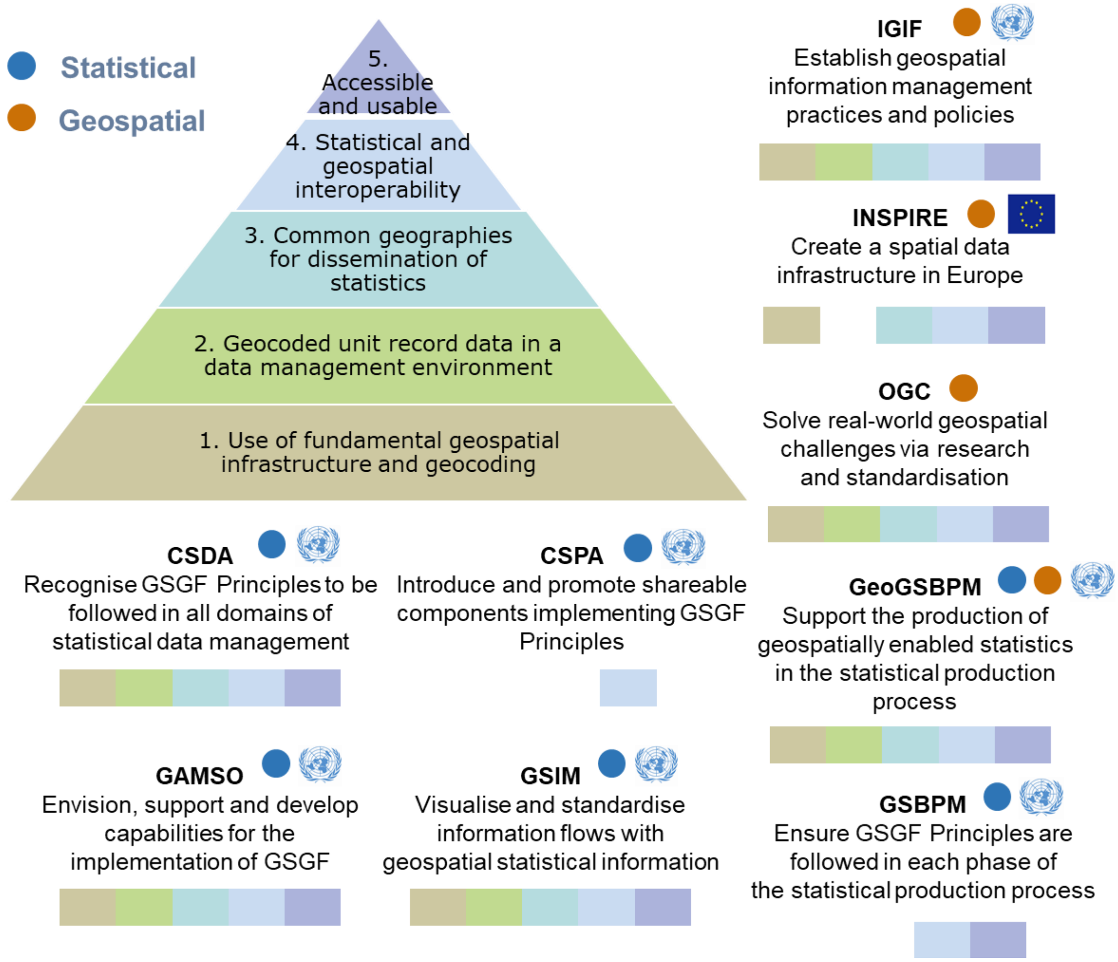

Figure 9. Nine frameworks and GSGF in the statistical and geospatial domains.

First, a rough division between the frameworks can be made: GSBPM, GAMSO, GSIM, CSDA and CSPA have been developed and peer reviewed by the international statistical community, whereas the IGIF, INSPIRE, and OGC have a geospatial focus. GeoGSBPM was a combined development.

The frameworks can often be viewed as a "cookbook" for ensuring that the necessary aspects are identified, e.g. within all statistical business process steps (GSBPM) or information objects (GSIM). Generally, the frameworks also aim to provide a common vocabulary and support collaboration activities, particularly in the field of modernisation. The frameworks can thus support the implementation of the GSGF within European context.

The frameworks are not developed in isolation but must be seen in the context of a whole suite of standards and frameworks, e.g. CSDA links to GSBPM, GSIM and CSPA and greater value can be obtained from the GSIM if it is applied in conjunction with the GSBPM. Different frameworks support the work from different aspects, and they complement each other but should be used differently.

5.1 GSBPM and GeoGSBPM: Identify Activities and Tasks to Convert Data into Statistical Information

The Generic Statistical Business Process Model, GSBPM (UNECE, 2019a), is a collection of related and structured activities and tasks for converting input data into statistical information, i.e. the statistical business process. Furthermore, the GSBPM recognises several overarching processes with a strong statistical component that apply throughout the eight phases of the statistical business process: quality management; metadata management; data management; process data management; knowledge management; and provider management.

One HLG-MOS task team has produced a document describing geospatial related activities under each sub-process of the GSBPM. The results of the task team with the GSBPM are a good tool for describing the geospatially enabled statistical business process, the “GeoGSBPM” (UNECE, 2021a). It provides a steady starting point to describe the interactions between the statistical and geospatial pipelines, thereby building consistent flows between these two pipelines according to GSGF principles. The results of this HLG-MOS task team have been used in the study of the GSGF Europe Reference Architecture.

5.2 GAMSO: Define Capabilities Needed to Support Statistical Production

The Generic Activity Model for Statistical Organisations, GAMSO (UNECE, 2019b), extends and complements the GSBPM by identifying additional activities needed to support statistical production. While individual collaboration typically focuses on modernising a particular aspect of production (as described by the GSBPM), statistical production occurs within a broader context of corporate strategies, capability development and corporate support which are the focus for GAMSO.

Changing how statistics work is a slow process, with many people involved. GAMSO focuses on widescale assets that make the overall implementation of the GSGF possible. For example, studying GAMSO reveals a need for more support from management and high-level strategic activities. Much cooperation between different pipelines is also required inside and between organisations for the GSGF to be implemented.

Furthermore, when the GSGF is studied from the perspective of GAMSO, capability development supports the implementation of the GSGF by providing a common approach to help compare an organisation’s own capabilities to those required, and to identify and plan capability improvements. In addition to the development of capability elements, there is also a challenge of maintenance over time. In the implementation of the GSGF, GAMSO can thus be used to examine long-term goals, transformation and cultural questions such as how organisations may become more used to using geospatial data to enhance statistics.

In Europe, the implementation of the GSGF is much more about policies, responsibilities, resources, communication, collaboration and commitment to common vision and strategic goals than, for example, the availability of IT systems and technology. Thus, paying attention to all three activity areas identified in GAMSO can support the work on the way to implementing the principles defined by GSGF.

5.3 GSIM: Define and Describe Information Objects in a Harmonized Way

The Generic Statistical Information Model, the GSIM (UNECE, 2019c), is a reference framework to define and describe the pieces of information used in the production of official statistics. The GSIM also enables generic descriptions of both the management and use of data and metadata throughout the statistical production process.

The GSBPM and the GSIM are recognised as key standards for the modernisation of official statistics. They aim to provide common terminology that improves communication about the production of statistics within and between organisations. Identifying links between the GSIM and the GSGF can produce a combined set of metadata, thereby supporting the implementation of the GSGF. Furthermore, the GSIM could support the visualisation and standardisation of information flows with geospatial information. This may facilitate collaboration and the reuse and sharing of good practices, methods, components and processes, thereby leading to greater efficiency.

5.4 CSDA: Identify Capabilities Needed to Manage Data

Common Statistical Data Architecture, CSDA (UNECE, 2018), shows organisations how to organise and structure their processes and systems for the efficient and effective management of data and metadata. The scope of CSDA includes all the GSBPM phases. Besides the operational phases, CSDA addresses cross-cutting issues: data governance; traceability; quality and security. Non-statistical business processes such as HR and Finance are beyond the scope of CSDA.

There is no restriction to the types of data in CSDA. In particular, to help organisations modernise themselves, CSDA shows how to deal with the newer types of data sources such as big data, scanner data and web scraping. Furthermore, there is no data without metadata. They are both considered integral information. Although CSDA is loosely based on TOGAF, it should be stressed that “data” for statistical organisations means something different from what is understood by most industries. “Data” for statistical organisations is the raw material, the parts and components and the finished products, rather than the information needed to support and execute the organisation’s primary processes. That said, some data play that role within the NSI, of course. Thus, CSDA is not what enterprise architecture traditionally refers to by “data architecture”.

CSDA provides statistical organisations with some key principles that aim to direct how capabilities are identified and defined. These key principles overlap and support those of the GSGF. Like the process phases in the GSBPM, CSDA can be used as a tool to recognise GSGF principles to be followed in each domain/capability of statistical data management. Furthermore, the CSDA capability model may help in designing business and application layers.

5.5 CSPA: Develop, Share and Reuse Components in Statistical Production

Common Statistical Production Architecture, CSPA (UNECE, 2021b), is an industry architecture for the official statistics industry, defined as a set of agreed common principles and definitions designed to promote greater interoperability within and between the different stakeholders comprising an industry. CSPA also aims to reduce the cost of developing and maintaining processes and systems.

CSPA builds on and uses GAMSO, the GSBPM and the GSIM to improve the common understanding and alignment necessary for the joint development, sharing and reuse of components. It gives users an understanding of the different statistical production elements (i.e. processes, information, applications, services) that make up a statistical organisation, and how these elements are relate. It also provides a common vocabulary with which to discuss implementations, with the aim of enhancing commonality. Furthermore, CSPA enables the vision and strategy of the statistical industry by providing a clear, cohesive and achievable picture of what is required to get there.

Within CSPA, the focus is on the application architecture layer. CSPA has an immediate connection with GSGF Principle 4. For statistical pipelines, CSPA provides a generic way to describe services implementation, based on interoperability and standardisation, as is done by INSPIRE on the geospatial side. Both CSPA and INSPIRE support standardised architecture. CSPA promotes shareable services using a common breakdown of supported statistical functionalities, whereas geospatial services are included in the INSPIRE catalogue.

5.6 IGIF: Develop and Strengthen National Geospatial Information

The Integrated Geospatial Information Framework, the IGIF (UN-GGIM, 2020b), is a strategic guide to enable the coordination, development, strengthening and promotion of effective sharing and use of geospatial information for policy formulation, decision making and innovation. The IGIF focuses on concepts, methods, standards and guides to address global goals, as well as the national needs of all countries. Structurally, the IGIF comprises a strategic-level overarching guide, a practical-level implementation guide focusing on nine strategic pathways and a country-level action plan.

NSDIs have traditionally focused on the management and dissemination of data using standardised interoperable web services. The emphasis has been on technological implementations of data harmonisation and services. In contrast, the IGIF additionally focuses on the governance, policy, financial, capacity and engagement processes necessary to collect, maintain, integrate and share geospatial information through all levels of government and society in a modern and enabling technology environment.

Where GAMSO promotes the supporting actions within an organisation, the IGIF is more useful at an interorganisational level and helps identify tasks that require a shared responsibility between organisations. Most of the IGIF strategic pathways coincide, align with and complement GSGF key elements and support implementation of the GSGF.

The GSGF mostly considers the integration of fundamental authoritative data sources from governmental data providers and custodians, with clearly agreed roles and responsibilities defined by legislation. Emerging new data domains – for example, health and non-authoritative data sources and new users such as those in community and citizens, where no formally agreed roles and responsibilities are defined – pose a challenge for NSI and NGIAs in GSGF implementation. Here, the IGIF provides guidance on actions, methodology, and tools to cover and capture the potential of non-authoritative data collection, management and delivery. However, the IGIF's global nature means a careful selection of the useful elements that can be utilised for GSGF Europe implementation must be carried out within each of the nine IGIF strategic pathways.

5.7 INSPIRE: Create a Common Spatial Data Infrastructure

The infrastructure for spatial information in Europe, the INSPIRE directive (EU, 2007), creates a European Union spatial data infrastructure for the purposes of EU environmental policies and policies or activities which may have an impact on the environment. INSPIRE aims to enable the sharing of environmental spatial information among public sector organisations, facilitate public access to spatial information across Europe and assist in policy-making across boundaries. It comprises five common principles and common Implementing Rules, and addresses 34 spatial data themes needed for environmental applications.

INSPIRE is recognised as an important framework under GSGF Europe. The main goal of INSPIRE is to allow the full discovery, accessibility and sharing of the national geospatial datasets owned by the public sector. INSPIRE provides important support for interoperability, but also lacks some aspects e.g. the versioning and linkages of historic data and developing geographies.

5.8 OGC: Improve Access to Geospatial or Location Information

The Open Geospatial Consortium (OGC) is a worldwide community committed to improving access to geospatial or location information. The OGC creates free and publicly available geospatial standards that enable new technologies. The OGC also manages an agile and collaborative research and development process – the OGC Innovation Program – that anticipates and solves the real-world geospatial challenges experienced by OGC members. The mission of the OGC is to make location information more Findable, Accessible, Interoperable, and Reusable (FAIR, 2016).

The complexity of bringing together survey, administrative and big data with geospatial information and earth observation is driving an increased need of standardisation to support better statistical outputs. The OGC Statistical Domain Working Group was established on the premise that the statistical domain was evolving away from the traditional capture of data through census and other surveys to a more dynamic integration of data from across a wide range of domains that could make statistics more timely, more accurate and better understood. It states that geospatial information is a key data source for transforming how statistics are produced, because it has the potential to fit into every component of the Generic Statistical Business Process Model (GSBPM).

For the GSGF, the OGC is mainly a standards and technological framework that is aware of the potential that can be achieved by the integration of geospatial and statistical information or through the creation of geospatial statistics.

OGC activities, standards and technologies support the goals and recommendations of the GSGF by providing interoperability support across all five GSGF principles.

6. Terms and Definitions

Accessibility - Accessibility together with clarity refers to the conditions and modalities by which users can obtain, use and interpret data.

Accuracy - Accuracy refers to the closeness of estimates to the unknown true values.

Big data - Big data is an all-encompassing term for any collection of data sets so large or complex that it becomes difficult to process them using traditional data-processing applications.

Capability - Capability means something that enables us to carry out an activity: An ability that an organisation, person or system possesses. Capabilities are typically expressed in general and high-level terms and typically require a combination of organisation, people, processes and technology to achieve (Source: The Open Group).

Clarity – Clarity together with accessibility refers to the conditions and modalities by which users can obtain, use and interpret data.

Coherence - Coherence refers to the adequacy of the data to be reliably combined in different ways and for various uses.

Comparability - Comparability refers to the measurement of the impact of differences in applied statistical concepts, measurement tools and procedures where statistics are compared between geographical areas, sectoral domains or over time.

Crowdsourcing - Crowdsourcing is "the practice of obtaining needed services, ideas or content by soliciting contributions from a large group of people and especially from the online community rather than from traditional employees or suppliers". (Source: Merriam-Webster)

Data ecosystem - In this context, a data ecosystem refers to a network of organisations involved in the delivery of data products and/or services through cooperation. A data ecosystem combines data from numerous providers and builds value through the usage of processed data e.g., the statistical-geospatial (data) ecosystem.

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. (Source: The GSGF global)

Enterprise architecture, EA - Enterprise Architecture is about understanding all of the different elements involved in making up an enterprise and how those elements interrelate. It is an approach to enabling the vision and strategy of an organisation, by providing a clear, cohesive and achievable picture of what is required to get there. (Source: Statistical Network BA definition)

Framework - A framework is a structure for a content or process that can be used as a tool to structure thinking, ensuring consistency and completeness (TOGAF).

Geocoding - Geocoding is the process of transforming a description of location or unreferenced location information (such as address, a name of a place or other location information) to a location's measurable position on the earth's surface. In other words, geocoding is a way to make data geospatially enable to indicate where the data are in space by linking unreferenced location information associated with a feature through a unique identifier (e.g., housing unit or business) to a set of coordinates within a coordinate system (also referred to as a spatial reference system). These resulting coordinates are the geocode.

Geography – Geography represents a division of space into smaller units, typically areas.

Georeferencing (also geospatial referencing) – Georeferencing is generally defined as the process of associating statistical unit records (or other non-spatial data) to a specific location in space for use in geospatial analysis and consists of a set of broad processes that include geocoding. Georeferencing 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 analysed with other geographic data. Georeferencing can be defined shortly as "relating data with where the fact happens" (Source: GSBPM).

Geospatial information – Generally, geospatial information is information that has traditionally been portrayed through maps or in association with maps. More technically, geospatial information is defined as information with a direct or indirect reference to a specific location or geographical area on or near the surface of the earth. Geospatial information generally relates to the natural and built environment, but also includes observations of people, and the social and economic outcomes of human activity. Geospatial information is stored in a geographic referencing system, usually a coordinate system of longitude, latitude and, increasingly, elevation.

Geospatial object - When a statistical unit is linked to location, it becomes a geospatial object which has a direct link to geospatial terminology and geospatial elements. Geospatial object means an abstract representation of a real-world phenomenon related to a specific location or geographical area. Geospatial objects are represented with their geometry (point, line, polygon).

Geospatial reference data - The geospatial reference data contain preferably authoritative geospatial information to be used to geocode directly or indirectly all public sector information at all levels of government, including data sources for statistical information.

Geospatial statistics (also geospatially enabled statistics) - Location or geospatial-extent are the main characteristics of geospatial statistics. Spatial statistics are geocoded to small (in most cases below level 5) administrative or non-administrative geographies. Spatial statistics may also result from the integration of statistical and geospatial information during the statistical production process, although the product might be regional statistics. The cross-border of functional perspective might be another important factor to define statistics as geospatial.

Grid - Grid is network composed of two or more sets of curves in which the members of each set intersect the members of the other sets in an algorithmic way. (Source: ISO 19123:2005)

Interoperability – Interoperability is a capability to communicate, execute programs or transfer data among various functional units in a manner that requires the user to have little or no knowledge of the unique characteristics of those units. (Source: ISO/IEC 2382:2009, 2121317)

Metadata - Metadata are data about data that define and describe and provide information about other data.

Microdata - Microdata are data on an individual object obtained, e.g. from sample surveys, censuses and administrative systems.

(National) Spatial Data Infrastructure, (N)SDI (also geospatial infrastructure) – A (National) Spatial Data Infrastructure is the technology, policies, standards, good practices and human resources necessary to acquire, process, store, distribute and improve utilisation of geospatial data. A successful (N)SDI implementation addresses the following considerations:

- Maintenance of data and systems,
- Redundancies should be built into the dissemination solution to prevent a single point of failure,
- Final review and pre-processing before release (data disclosure and confidentiality) to prevent disclosure issues, and
- Generalisation and thinning of spatial data should be implemented to ensure that the data meet the minimum level of quality and are useable at defined scaled supporting both large and small-scale needs. This can impact both cartographic and data storage issues.

Open data - The data to be freely accessed, used, modified, integrated and shared for any purpose.

Open data licence - Open data licence is an agreement to allow users to publish, provide and use data freely as open data.

Open standard - Open Standards are standards made available to the general public and are developed (or approved) and maintained via a collaborative and consensus driven process. Open Standards facilitate interoperability and data exchange among different products or services and are intended for widespread adoption. (Source: International Telecommunication Union, ITU)

Operating environment - Here, an operating environment is a sum or a collection of all internal and external factors, which have a direct or indirect bearing on the functioning of organisations, such as economic, social, political, legal and technical factors.

Punctuality - Punctuality refers to the possible time lag existing between the actual delivery date of data and the target date when it should have been delivered, for instance, with reference to dates announced in some official release calendar or previously agreed among partners. (Source: OECD Glossary of Statistical Terms)

Reference architecture - A reference architecture is a generic architecture that provides guidelines and options for making decisions in the development of more specific architectures and the implementation of solutions. A reference architecture can be at any point of the architecture continuum. (Source: The Open Group)

Relevance – Relevance refers to the degree to which data and information meet current and potential needs of the users.

Representativeness - A term used to describe how characteristic a particular item is of the types of goods and services included in a basic heading. (Source: OECD glossary of statistical terms)

Scanner data - Scanner data are the data recorded by the retailers when consumers make purchases. They include, for each article sold in a store on a given day, the quantity and the sales price of the articles sold. (Source: Insee)

Timeliness - Timeliness refers to the period between the availability of the information and the event or phenomenon it describes.

Web scraping - Web scraping is a technique to extract data from web sites. Typically, automated processes are implemented but web scraping can also be done manually.

7. References

Brady, M. (2015). Towards a Global Statistical Geospatial Framework. Australian Bureau of Statistics, keynote presentation EFGS 2015 Conference Vienna, Austria. (https://www.efgs.info/wp-content/uploads/conferences/efgs/2015/Conference_EFGS2015_1010_0_ABS_Martin-BradyPresentation.pptx)

Calder, A. (2021). Geospatial ethics for statistics. ONS Office for National Statistics, UK. Presentation EFGS 2021 Virtual Conference. (https://www.efgs2021.de/EFGS2021/programme/presentation/session2-5_calder.pdf?__blob=publicationFile)

EC (2017). The New European Interoperability Framework. A website. (https://ec.europa.eu/isa2/eif_en)

EC, DG REGIO (2018). A walk to the park? Assessing access to green areas in Europe's cities. Working papers, 01/2018. (http://ec.europa.eu/regional_policy/en/information/publications/working-papers/2018/a-walk-to-the-park-assessing-access-to-green-areas-in-europe-s-cities)

EFGS (2021). The European Forum for Geography and Statistics. A website. (https://www.efgs.info)

ESS (2019). Quality Assurance Framework of the European Statistical System. Version 2.0. (https://ec.europa.eu/eurostat/documents/64157/4392716/ESS-QAF-V1-2final.pdf/bbf5970c-1adf-46c8-afc3-58ce177a0646)

EU (2021a). Europe's Digital Decade: digital targets for 2030. (https://ec.europa.eu/info/strategy/priorities-2019-2024/europe-fit-digital-age/europes-digital-decade-digital-targets-2030_en)

EU (2021b). EU Data Strategy 2020. (http://www.dataeconomy.eu/eu-data-strategy-2020/)

EU (2019). European Green Deal. (https://ec.europa.eu/info/strategy/priorities-2019-2024/european-green-deal_en)

EU (2016). General Data Protection Regulation. (https://eur-lex.europa.eu/eli/reg/2016/679/oj)

EU (2007). Infrastructure for spatial information in Europe. INSPIRE. EU Directive. (https://inspire.ec.europa.eu/)

Eurostat (2021a). Data Protection – Rules for the protection of personal data inside and outside the EU. (https://ec.europa.eu/info/law/law-topic/data-protection_en)

Eurostat (2021b). Statistical Confidentiality. (https://ec.europa.eu/eurostat/web/research-methodology/statistical-confidentiality.)

Eurostat (2020). European harmonised Labour Market Areas — methodology on functional geographies with potential. Statistical working papers. ISBN 978-92-76-15370-2 (https://ec.europa.eu/eurostat/documents/3888793/10992115/KS-TC-20-002-EN-N.pdf/5a6258be-6c53-64fa-971c-2b19908ab442)

Eurostat (2019). The integration of statistical and geospatial information – a call for political action in Europe. (https://un-ggim-europe.org/wp-content/uploads/2019/09/KS0319423ENC_new.pdf)

Eurostat (2018). Methodological manual on territorial typologies – 2018 edition. (https://ec.europa.eu/eurostat/web/products-manuals-and-guidelines/-/KS-GQ-18-008)

FAIR (2016). FAIR Principles. (https://www.go-fair.org/fair-principles/)

GEOSTAT 1A (2012). GEOSTAT 1A – Representing Census data in a European population grid. GEOSTAT 1A project. EFGS and Eurostat. (https://www.efgs.info/wp-content/uploads/geostat/1a/GEOSTAT1A-final-report.pdf)

GEOSTAT 3 (2019). GSGF Europe – Implementation guide for the Global Statistical Geospatial Framework in Europe. Proposal from the GEOSTAT 3 project. EFGS and Eurostat. (https://www.efgs.info/wp-content/uploads/geostat/3/GEOSTAT3_GSGF_EuropeanImplementationGuide_v1.0.pdf)

GEOSTAT 4 (2021a). GSGF Europe: Reference Architecture. A stand-alone paper produced by GEOSTAT 4 project. GEOSTAT 4 and Eurostat.

GEOSTAT 4 (2021b). GSGF Europe: Requirements and Recommendations. A stand-alone paper produced by GEOSTAT 4 project. GEOSTAT 4 and Eurostat.

OGC (2021). Open Geospatial Consortium. (https://www.ogc.org/)

Tobler, Waldo R. (1970). "A Computer Movie Simulating Urban Growth in the Detroit Region." Economic Geography (Supplement: Proceedings, International Geographical Union. Commission on Quantitative Methods), 46: 234–240. DOI:10.2307/143141

UK Statistics Authority (2021). Ethical considerations in the use of geospatial data for research and statistics. A guidance document. (https://uksa.statisticsauthority.gov.uk/publication/ethical-considerations-in-the-use-of-geospatial-data-for-research-and-statistics/pages/1/)

UN (2015). The 2030 Agenda for Sustainable Development. (https://sdgs.un.org/goals)

UNECE (2021a). Geospatial view of Generic Statistical Business Process Model. GeoGSBPM. Version 1.0. (https://statswiki.unece.org/display/GSBPM/GeoGSBPM)

UNECE (2021b). Common Statistical Production Architecture. CSPA. Version 2.0. (https://statswiki.unece.org/display/CSPA/CSPA+v2.0)

UNECE (2019a). Generic Statistical Business Process Model. GSBPM. Version 5.1. (https://statswiki.unece.org/display/GSBPM/GSBPM+v5.1)

UNECE (2019b). Generic Activity Model for Statistical Organisations. GAMSO. Version 1.2. (https://statswiki.unece.org/display/GAMSO/GAMSO+v1.2)

UNECE (2019c). Generic Statistical Information Model. GSIM. Version 1.2. (https://statswiki.unece.org/display/gsim/Generic+Statistical+Information+Model)

UNECE (2018). Common Statistical Data Architecture. CSDA. Version 2.0. (https://statswiki.unece.org/display/DA/CSDA+2.0)

UN-GGIM (2021). The Global Statistical Geospatial Framework: Draft Implementation Guide. GGIM 11th Session. (https://ggim.un.org/meetings/GGIM-committee/11th-Session/documents/GSGF%20-%20Implementation%20Guidance.pdf)

UN-GGIM (2020a). The Integrated Geospatial Information Framework (IGIF) and the Global Statistical Geospatial Framework (GSGF). Greg Scott, UN-GGIM Secretariat, EFGS 2020 Webinar. (https://www.efgs.info/wp-content/uploads/2020/10/EFGS_2020_session2_20.10.2020_Greg-Scott.pdf)

UN-GGIM (2020b). Integrated Geospatial Information Framework. IGIF. (https://ggim.un.org/IGIF/overview/)

UN-GGIM (2015). A Guide to the Role of Standards in Geospatial Information Management. August 2015, Chapter 2. (https://ggim.un.org/documents/standards%20guide%20for%20unggim%20-%20final.pdf)

UNSC and UN-GGIM (2019). The Global Statistical Geospatial Framework. The GSGF. (https://ggim.un.org/meetings/GGIM-committee/9th-Session/documents/The_GSGF.pdf)

8. Annexes

Annex I – Requirements and Recommendations

A set of requirements (Figure I.a) and recommendations connected to these requirements were identified and described to support the implementation of the GSGF in Europe. These requirements and recommendations are presented in the GSGF Europe: Requirements and Recommendations document (GEOSTAT 4 and Eurostat, 2021b).

Requirements	Principle
5.1 Implement clear and simple data licensing policies 5.2 Use service- oriented data portals supporting dynamic integration of data 5.3 Define clear national and European rules to ensure protection of privacy 5.4 Facilitate data search and use through cataloguing and improved guidance	P5: Accessible and usable geospatially enabled statistics
4.1 Improve geospatial workflows within statistical production 4.2 Enable data integration through consistent semantics and concepts across domains 4.3 Publish data once and leave them at their source to be reused many times 4.4 Increase use of services and semantic web technology to enable innovation in a wider data ecosystem	P4: Statistical and geospatial interoperability
3.1 Set up and maintain a consistent framework of national statistical and administrative geographies 3.2 Consolidate use of existing statistical grid systems and explore the potential of evolving global grid systems	P3: Common geographies for the dissemination of statistics
2.1 Build an effective and secure data management environment 2.2 Include a geospatial aspect in organisation's enterprise architecture 2.3 Ensure consistency and quality of geocoding results 2.4 Use point-of-entry validation in collection of administrative or statistical data 2.5 Define common data quality frameworks taking into account spatial and temporal consistency	P2: Geocoded unit record data in a data management environment
1.1 Use data from National Spatial Data Infrastructures. 1.2 Use point-based reference data for geocoding. 1.3 Build formal working relationships on institutional agreements.	P1: Use of fundamental geospatial infrastructure and geocoding

Figure Ia. GEOSTAT Europe: Requirements to support implementation of the GSGF Europe.

Annex II – Some Examples of Major Standards

Data format

- Data format standards are designed to define technical specifications for organising and storing geospatial data in a given format. For example, GeoPackage, defined by the Open Geospatial Consortium (OGC), is an open, non-proprietary, platform-independent and standard-based data format for geographic information systems. Keyhole Markup Language is another example of data format standards for expressing geographic annotation and visualisation within two-dimensional maps and three-dimensional Earth browsers. Keyhole Markup Language was developed for use with Google Earth. It then became an international standard of the Open Geospatial Consortium.

Metadata

- ISO 19115 is a metadata standard which defines how to describe geographical information and services by means of metadata.

Discovery and access

- GeoDCAT-AP is an extension of the Data Catalogue Application Profile (DCAT-AP) for describing geospatial datasets, dataset series, and services. Its basic use case is to make spatial datasets, data series, and services searchable on general data portals, thereby making geospatial information more discoverable across borders and sectors. It provides a syntax using metadata elements as defined in the framework of the INSPIRE directive.
- The OGC API Standard provides the fundamental API building blocks to create, modify and query "features" on the Web – enabling a much simpler way to share and access geospatial information that is consistent with the architecture of the Web.

Services

- The OGC Table Joining Service standard defines an interface for services that provide the ability to join attribute data stored in one database on a network with corresponding geometry (points, lines or polygons) stored in another network accessible database.
- the OpenGIS Web Map Service standard defines an interface that allows a client to obtain maps of geospatial data and gain detailed information on specific features shown on the map.
- The Web Feature Service standard is for sharing feature data across the web for direct access to geospatial data.

Annex III – The Four GEOSTAT Projects

The need to enhance the use of the spatial dimension during the collection, processing and dissemination of statistics was acknowledged by the EU and its Member States about two decades ago. Drawing on the efforts made by the Nordic countries to develop harmonised gridded population data in the 1990s, Eurostat initiated the "TANDEM project" in 1999. Since 2010, Eurostat has supported ESSnet GEOSTAT projects with the goal of building a methodological foundation for the integration of statistical and geospatial information and for the production of geospatial statistics in the European Statistical System (ESS).

The approach of the suite of GEOSTAT projects was to start with a practical use case as proof of concept, and then derive general guidelines and methodology applicable to other statistical domains. Each GEOSTAT project built on the results of the previous project and expanded the scope and ambition. All the materials produced by the GEOSTAT projects is available online at https://www.efgs.info/geostat/.

The contributors of each GEOSTAT project:

- GEOSTAT 1 "Representing census data in a European population grid dataset"
  - Statistics Norway (coordinator of part 1A), NSIs from Austria, Estonia, Finland, France, the Netherlands, Poland, Portugal and Slovenia and 1 sub-contractor MD Mapping.
  - National Statistical Institute of Bulgaria (coordinator of part 1B), NSIs from Czechia, Estonia, Finland, Norway and Portugal, and 1 sub-contractor MD Mapping.
- GEOSTAT 2 "A point-based foundation for statistics"
  - Statistics Sweden (coordinator), NSIs from Austria, Finland, France, Norway, Poland and Portugal
- GEOSTAT 3 "A statistical geospatial framework for sustainable development"
  - Statistics Sweden (coordinator), NSIs from Portugal, Norway, Austria, Estonia, Finland, Poland and the Netherlands
- GEOSTAT 4 "Geospatial statistics"
  - Statistics Finland (coordinator), NSIs from Sweden, Austria, Poland, Germany, Slovenia, France, Norway and Portugal, and sub-contractors National Land Survey of Finland, Kartverket (The Norwegian Mapping Authority) and MD Mapping.