GSGF Finland (GSFI)

Title: GSGF Finland (GSFI)
Project: The GSGF in Finland — Integration of geospatial and statistical information in Finland (GSFI)
Grant agreement number: 101112903 - 2022-FI-GEOS-GSFI

Content

1. Introduction

One of the central goals of the public administration strategy in Finland is that preparation and decision-making are guided by information (Ministry of Finance, 2020). The efficient sharing and use of information in a complex world requires interoperability from the data produced by different actors. The easy integration and usability of data in different applications also supports innovation, the development of new technologies and services.

The integration of statistics and geospatial data is a crucial component of modern data policy, aiming to leverage diverse data sources to support decision-making. Statistics provide quantitative information and trends, while geospatial data adds a geographical dimension, helping to understand regional differences and local specificities. Combining these data types enables deeper analysis and visualisations that make complex information more comprehensible and accessible. High-quality geospatial information also plays a significant role in interoperability, as it enables the sharing of comparable information between different systems and organisations. Time and money are saved when there is less need for manual data entry and corrections. Integration of statistical and geospatial data is particularly important in complex fields such as environmental protection, urban planning, and healthcare.

A high-level framework, the Global Statistical Geospatial Framework (GSGF), was published in 2019 to support the integration of statistics and geospatial data and to develop consistent production methods (UNSC and UN-GGIM, 2019). The European adaptation, GSGF Europe, was released in 2021 (GEOSTAT 4 and Eurostat 2021). It was recognised that examining global frameworks in relation to local conditions is crucial, as each country or region has its unique characteristics. For instance, local infrastructure, legislation, policies, user needs and resources may vary significantly. While the GSGF provides essential guidelines, only adapting it to local conditions ensures practical and efficient implementation of the framework, leading to the desired outcomes.

After the European adaptation, one more step in the development towards the concrete implementation of the reference framework was found necessary in Finland. The GSGF was interpreted into the national operating environment and GSGF Finland (GSFI) described in this document was created. In this work, the materials and instructions provided by GSGF Europe were of great help.

GSFI aims to guide national joint development by providing a common vision of national priorities and the roadmap to implement GSGF in Finland. Guidance is not obtained only to improve organisations own data management and production processes, but to understand better the needs of the operating environment and the production of stakeholders as well as to have a common view of the development needs of the geospatial data in the national registers.

In Chapter 2, the GSFI is presented, including:

- The path from the global GSGF to GSFI via GSGF Europe.
- GSFI vision and benefits.
- Principles of implementation describing the foundational shared guidelines for implementing GSFI effectively.
- Development roadmap, including objectives, measures, and a model for actor-specific development paths.

- A draft of the maintenance model supporting GSFI's ongoing maintenance, development, and expansion.
- Tools for practical work including a statistical geospatial glossary, GSFI reference architecture descriptions, a proposal for small area division, an analysis of confidentiality aspects in GSFI, and a study on value chains and data quality.
- Core Communication Materials including key communication resources, such as PowerPoint presentations both in Finnish and English and templates available in Finnish, Swedish and English.
- A library that initially contains a case study and present status review 2023.

The National Network for the Integration of Statistics and Geospatial Information in Finland (NISGIF) is an important stakeholder in the implementation of GSFI. The initiative to build GSFI came from the network itself. GSFI was constructed in collaboration with the network, and it is hoped that this will also enhance the cooperation of the organisations involved in the implementation of GSFI in the future. To strengthen the operation and efficiency of the network, the current state of the network was analysed. Experiences from building GSFI in cooperation with the network were also gathered. The goal was to identify development needs and proposals that foster cooperation within the network. The results are summarized in Chapter 3.

GSGF Finland (GSFI) was developed in the project "The GSGF in Finland — Integration of geospatial and statistical information in Finland (GSFI)". The project consortium included Statistics Finland, the Finnish Environment Institute and the National Land Survey of Finland. The project was financed by EU. The project was executed in close collaboration with the network NISGIF and through communication with Nordic and European colleagues. The results have been compiled in this document, which serves as a continuum for the global GSGF framework and GSGF Europe.

2. GSGF Finland

GSGF Finland (GSFI) is a national application of the Global Statistical Geospatial Framework (GSGF) (UNSC and UN-GGIM, 2019). GSFI comprises of four sections (Figure 1):

1. Strategic goals: The strategic goals presents an overview of the entity of the three levels of examination, including the GSGF, GSGF Europe and GSFI, vision for the work, and description of the benefits of the implementation.
2. Collaboration: The collaboration tells how the work itself progresses including principles of the implementation, the national roadmap and the GSFI maintenance model.
3. Tools: The first tools to support the work are the statistical geospatial glossary, GSFI reference architecture, a proposal for a small area division model, studies of confidentiality aspects and the use of value chains together with national data quality framework, and some material for communication.
4. Library: A case study of the business data production process in Statistics Finland and present status review in 2023 can be found in the library.

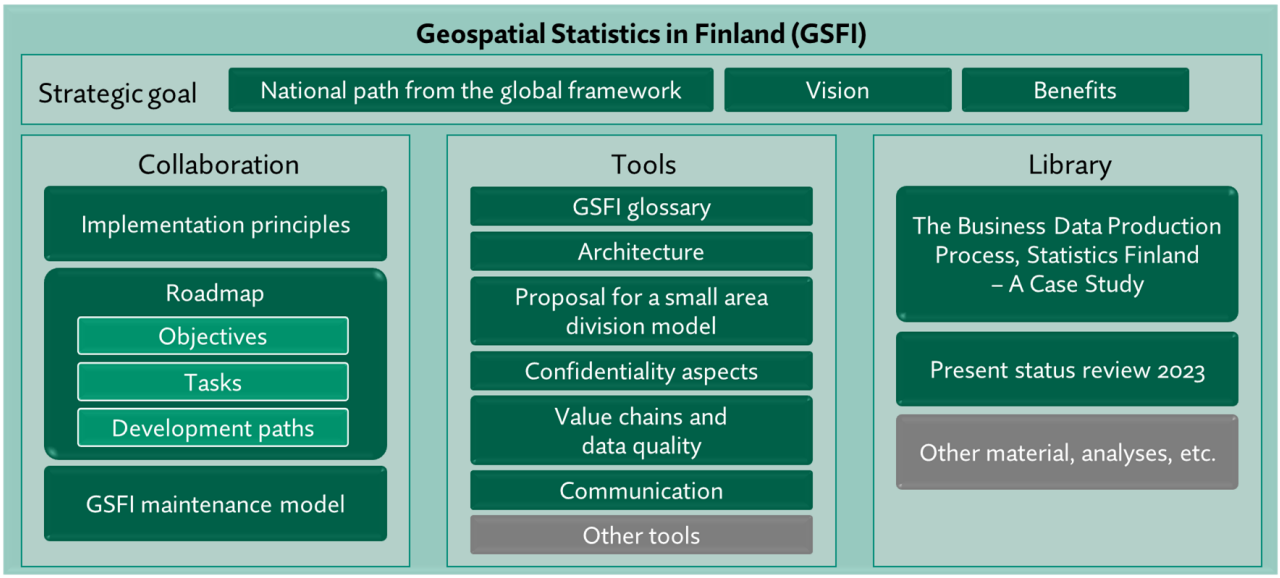

Figure 1. The structure of GSFI, version 1.0.

This document brings together the accumulated knowledge. More extensive ideas and observations are described in Finnish in separate documents that are available on the website "stat.fi/GSFI" and in the GSFI channel in Teams team of the network NISGIF, which is open to all network members. The aim is that the entity of GSFI is continuously maintained and developed so that the elements now available in GSFI will be supplemented.

2.1. National path from the global framework

For the national integration of statistics and geospatial information, there is available a global framework, the Global Statistical Geospatial Framework (GSGF). This framework provides a theoretical outline and perspective for understanding, structuring, and guiding concrete work. The first result in the implementation of the framework in Europe was providing the GSGF Europe information service. Now, this chain is complemented by a national adaptation of the GSGF in Finland (GSFI). Together, these three levels provide the model and tools for the development of the integration of statistical and geospatial information in Finland.

2.1.1. The Global Statistical Geospatial Framework (GSGF)

The Global Statistical Geospatial Framework (GSGF) provides the statistical and geospatial communities with a common high-level framework. It aims to connect socio-economic and environmental data to appropriate locations, improving the accessibility and usability of geospatial statistics. By applying its five principles and four key elements that support the implementation of these principles, the GSGF facilitates the production of harmonized and standardized geospatially enabled statistical data (Figure 2). The framework is available on the internet as a pdf document (UNSC and UN-GGIM, 2019).

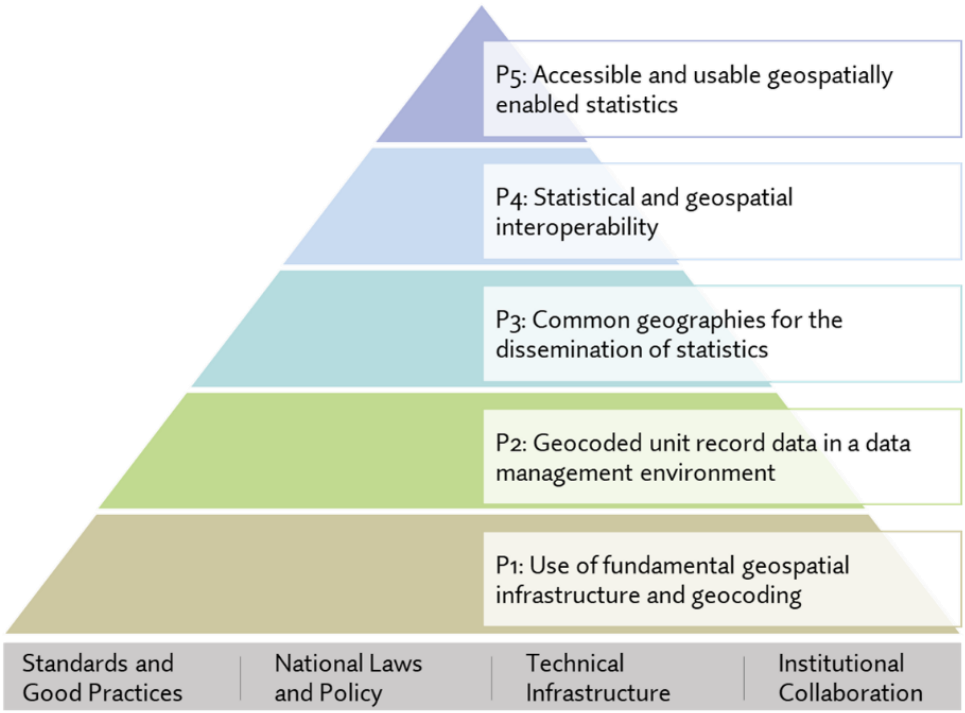

Figure 2. The five principles and four key elements of the GSGF (UNSC and UN-GGIM, 2019).

2.1.2. GSGF Europe

GSGF Europe builds a methodological foundation for the production of geospatial statistics in the European Statistical System by bringing the GSGF into the European context. It does not replace the descriptions and guidance provided at the global level but moves towards the framework's regional implementation. The five GSGF principles link to, and are implemented in, many levels within the European statistical and geospatial operating environment.

The structure and operation of the statistical geospatial community and the interpretation of the GSGF key elements are examined and four other aspects, i.e. quality, data collection, confidentiality, and innovation, were identified as essential for further study. Furthermore, the most relevant integration, standardisation and data-sharing frameworks and their links to the GSGF are also considered in GSGF Europe. All these aspects together are used to describe and analyse the European operating environment. This material is available on the GSGF Europe: GEOSTAT information service on the European Forum for Geography and Statistics (EFGS) website (EFGS, 2024).

2.1.3. National interpretation, GSGF Finland

The national adaptation of the GSGF, GSGF Finland (GSFI) is an interpretation of the global GSGF framework and GSGF Europe for the national operating environment. The aim of GSFI is to describe what the implementation of the framework would mean in practice in Finland. GSFI aims also to support progress towards this target state. The objectives and development needs are based on the comparison of the present status in Finland to the principles and analysis provided by the GSGF and GSGF Europe. This work laid the groundwork for the building of the GSFI road map. Furthermore, GSFI principles of implementation are based on the four key elements of the GSGF. In addition, the adaptation is supplemented with studies on integrating the newly launched national data quality framework and developing privacy policies. Important for the successful collaboration will also be the

core concepts introduced in a new Statistical Geospatial Glossary in Finnish. One key practical implementation in GSFI is the proposal for a new small area division model to enhance the national interoperability of geospatial statistics. This includes a nationally agreed method for producing hierarchical small areas for various thematic needs.

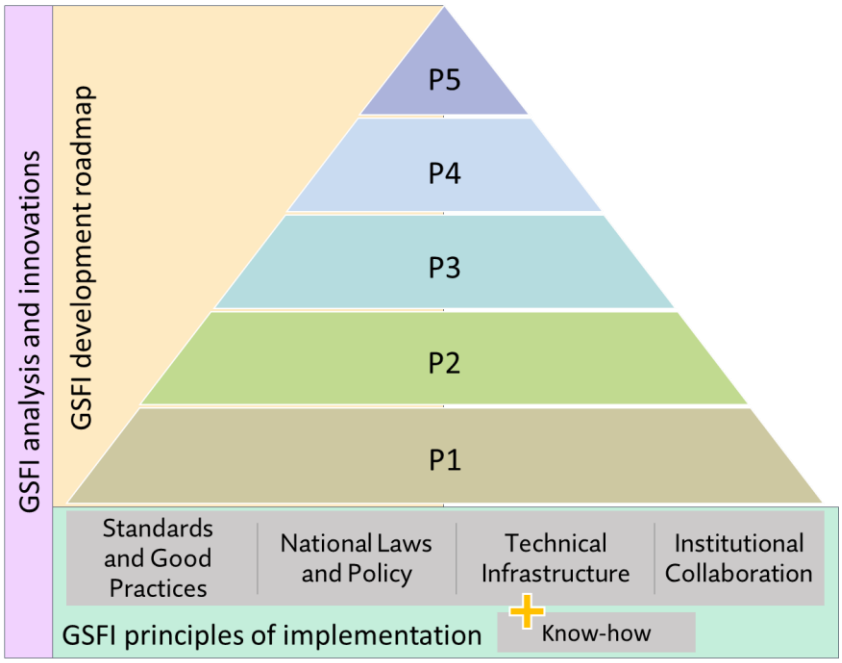

Figure 3. GSFI is being built around the global GSGF framework, expanding the GSGF principles towards practical implementation. The key elements of the GSGF have been interpreted as the implementation principles in GSFI. The national review has broadened the key element of technical infrastructure towards capability building, also taking into account know-how.

The successful implementation of GSFI relies strongly on national capability building, collaboration, and motivation among statistical, geospatial, administrative, and other data communities. The aim is that organisations share the GSFI vision and objectives and thereby also responsibilities and tasks. It is also expected that GSFI supports identifying mutual processes and shows benefits for the joint development.

GSFI is a living model. Its practical implementation requires that it evolves and develops with emerging new needs, and that elements which are no longer relevant due to changed circumstances are removed.

The GSGF influenced the understanding of the need for collaboration. The broad scope of the GSGF highlighted the need for comprehensive collaboration. As a result, Statistics Finland and the National Land Survey of Finland initiated the establishment of The National Network for the Integration of Statistics and Geospatial Information in Finland (NISGIF). The call to the community went out widely across the public, private, and third sectors. The general cooperation goals of the community, developed in collaboration, formed the foundation for building GSFI. NISGIF has also served as the reference group when creating the GSFI, working on the ideas and assumptions generated in the project. It is also clear that the NISGIF community will be the first one to which the GSFI will be 'handed over' once the project's work is completed.

2.2. Vision

In the vision of GSFI (Figure 4), Finland has a national, widely used spatial data infrastructure that provides location data for information and statistical production needs. The location data is based on unit-level information, making integration with other data efficient. The foundation of regional information is the national small area system and statistical grids. Location and regional data are discoverable and can be safely utilized. The societal significance of integrating statistics and spatial data has been recognized across administrative boundaries. The GSFI vision is available in Appendix A.

Figure 4. The GSFI vision: Geospatial data for information and statistics.

2.3. Benefits

The benefits of GSGF Finland's measures are extensive for geospatial and statistical ecosystems. The most important aspects include increased national efficiency, improved interoperability, enhanced data discoverability and availability, user support, quality and data protection management, and the support of data-driven research and decision-making. The better clarity of work and decision-making also supports motivation and innovation. Notable is that all these aspects also bring savings both in money and in time. In the following, it is analysed for each GSGF principle, what kind of benefits could be expected from the implementation of GSFI.

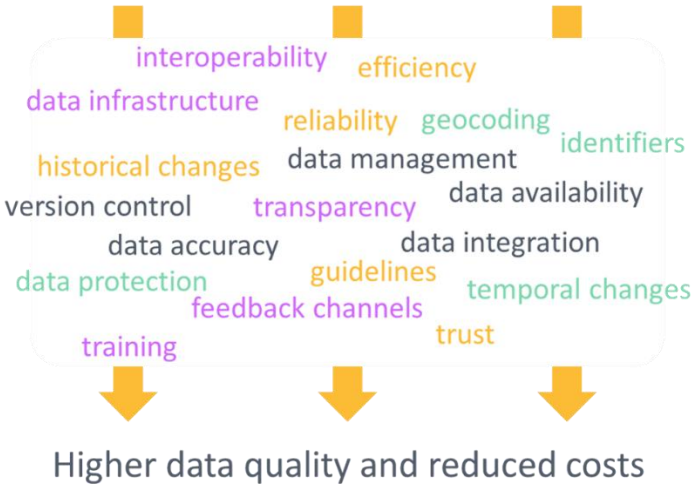

Figure 5. There are numerous expected benefits that can be summarised as improved data quality and reduced costs.

Principle 1. The efficiency, reliability, and interoperability of national data processes can be enhanced by defining key location data and responsible parties. Guidelines and development plans promote the consistent development of national spatial data infrastructure and effective data management. The national address information system improves data availability and accuracy for various needs. Uniform, point-based location data facilitate data integration, analysis, and enable more precise geocoding, which enhances data processes. Unique identifiers improve data interoperability and ensure reliable identification and linking of geospatial data objects. Up-to-date location data and different time versions allow for the tracking and analysis of historical changes. Clear guidelines and operating models improve the skills and efficiency of data producers and users, promoting broader use of national spatial data infrastructure. User-oriented training takes into account different skill levels, increasing the utilisation of geospatial data across various sectors. Efficient feedback channels and correction routines ensure that questions and issues are addressed promptly, improving the quality and reliability of the infrastructure.

Principle 2. By defining organisation-specific descriptions, development plans, and guidelines for the use of geospatial data, the utilisation of location information in organisational operations and development is improved, leading to more efficient processes. Preparing for future data infrastructure utilisation and its changes ensures that organisations can quickly and effectively leverage the national spatial data infrastructure, reducing transition issues and costs. Attaching location information and timestamps to data units improves data accuracy and interoperability, enabling better analysis and facilitating the integration of data from different sources. Attaching area identifiers and timestamps allows for flexible data aggregation and provides opportunities to better utilise data in geospatial analysis. Checking the quality of geospatial data at the data collection phase improves data quality and reliability throughout the information process. This reduces the number of errors and improves data usability. Considering the geospatial data perspective in data management ensures the consistency of data repositories and the effective management of temporal changes, which enhances data usability. Created operating models and descriptions of geospatial data sources improve data transparency and predictability, trust, and facilitate the use of data in various contexts. Agreed correction routines enable quick response and error correction, improving data quality and reliability.

Principle 3. Common national area divisions and classifications improve data interoperability, enabling more efficient integration and utilisation of data produced by different organisations. This reduces redundant work, as organisations do not need to produce and maintain their own methods for producing small area divisions. Additionally, common solutions and identifiers facilitate version control and change tracking of area divisions, improving data timeliness and reliability. Open availability and transparency increase when national area divisions and their historical data are accessible to all through interfaces. Furthermore, systematic collection and preservation of historical data enable the use of past statistics and information in current decision-making and research.

Principle 4. Benefits can be seen in the development of the geospatial-statistical ecosystem. Improving the division of labour and eliminating redundancies enhance cooperation between organisations and the efficient use of resources. Clear roles and responsibilities help avoid confusion and ensure that each organisation knows its tasks. Understanding customer needs and improving communication help develop geospatial information that better meets the needs of geospatial-statistical ecosystem actors. Improving communication also increases transparency and trust among all parties. Common conceptual models and target architecture improve the compatibility of data from different organisations, facilitate data sharing and utilisation and support co-development. Developing a statistical metadata model ensures that geospatial data is easily describable, discoverable, and usable for various purposes. Building national service packages promotes broader cooperation and ensures that geospatial data services complement each other. Defining standards and methods creates uniform practices that improve data quality and reliability. A maturity assessment model helps organisations identify their strengths and areas for improvement, supporting continuous improvement and learning.

Principle 5. Measures to improve data protection and geospatial data distribution bring many significant benefits. Uniform data protection practices reduce risks and ensure that all organisations adhere to the same data protection standards, increasing trust in data distribution and use. Emphasising data protection in communication improves awareness of its importance and ensures that all parties understand and follow agreed data protection practices. Common solutions in data processing facilitate the integration of data from different sources, which can lead to better and more accurate analysis results. Improving the discoverability and quality of geospatial data also brings significant benefits. Well-organised and presented information helps users find the information they need easily. saving This saves time and resources and improves data usability. Developing search services ensures that geospatial-statistical data are easily accessible and discoverable, promoting broader use and utilisation of data. Communicating about quality increases users' trust in the data. Supporting the use of location information in services for integrating data from different sources enables more diverse and accurate analyses. Using simple and uniform data structures improves data usability and facilitates the utilisation of data in various applications and services.

2.4. Collaboration

A wide range of different actors are involved in the development of statistical and geospatial integration in Finland. The work should therefore be viewed as development within the national information ecosystem, where successful implementation requires strong collaboration between different stakeholders. This collaboration is based on clear principles that guide the activities and a shared roadmap that defines the steps towards a common vision. Additionally, a governance model is needed to create the structures and mechanisms through which collaboration and development can be effectively maintained.

2.4.1. Implementation principles

The implementation principles aim to support efficient working and to ensure high-quality progress within GSFI. It is also expected that shared principles foster better collaboration among diverse stakeholders in the ecosystem. Common principles also ensure that all stakeholders are aligned in their approach, leading to consistent and standardised practices. This helps in reducing misunderstandings.

The aspects described now in the implementation principles came up repeatedly in the analysis of the current state and development needs in Finland, when the implementation of different GSGF principles was reviewed. It was found that the highlighted common needs could be connected to the key elements of the GSGF (Figure 6). The implementation principles are described in more detail in Appendix B.

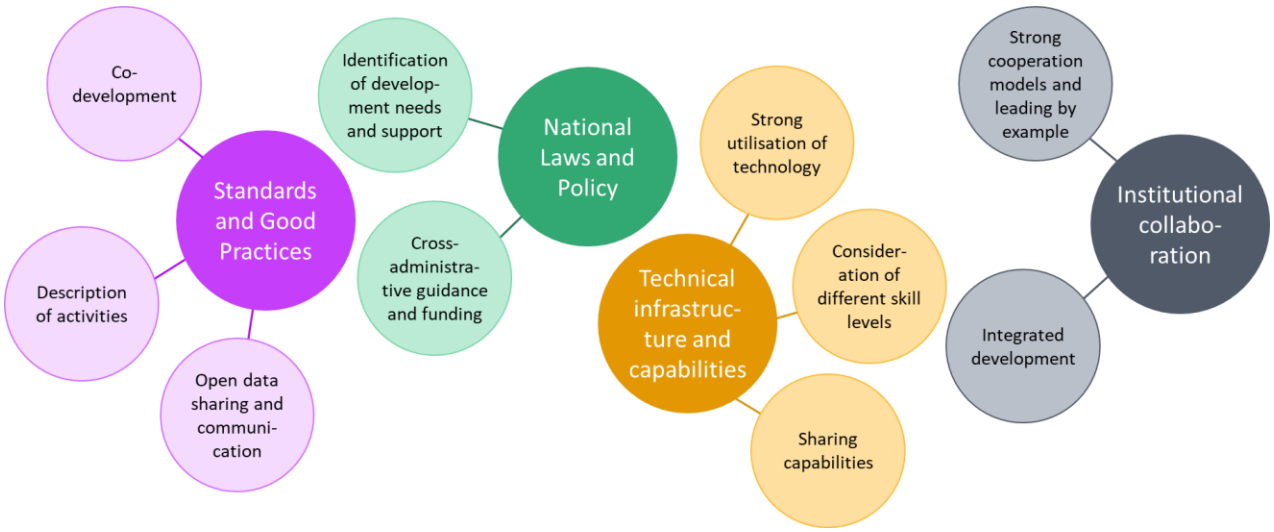

Figure 6. The GSFI principles of implementation are linked to the four key elements of the GSGF.

2.4.2. GSFI development roadmap

The GSFI development roadmap is a strategic plan that guides the progress of integrating statistics and geospatial information towards the GSFI vision. The goal is to provide a clear view of what is collectively aimed for and what is practically being done. In this type of ecosystemic co-development, the roadmap is particularly important because it helps ensure that everyone has the same situational awareness and that development proceeds as planned and efficiently. The roadmap also enables better communication and collaboration between different stakeholders, which is essential in a complex operational environment where actors have different perspectives, responsibilities, and expertise.

The roadmap is built on an understanding of the current situation, long-term objectives ensuring that all efforts are aligned with the vision, the tasks implementing the objectives, and the development paths of various stakeholders linked to the tasks. The roadmap is described in more detail in the document GSFI development roadmap, draft (GSFI and Eurostat, 2024a).

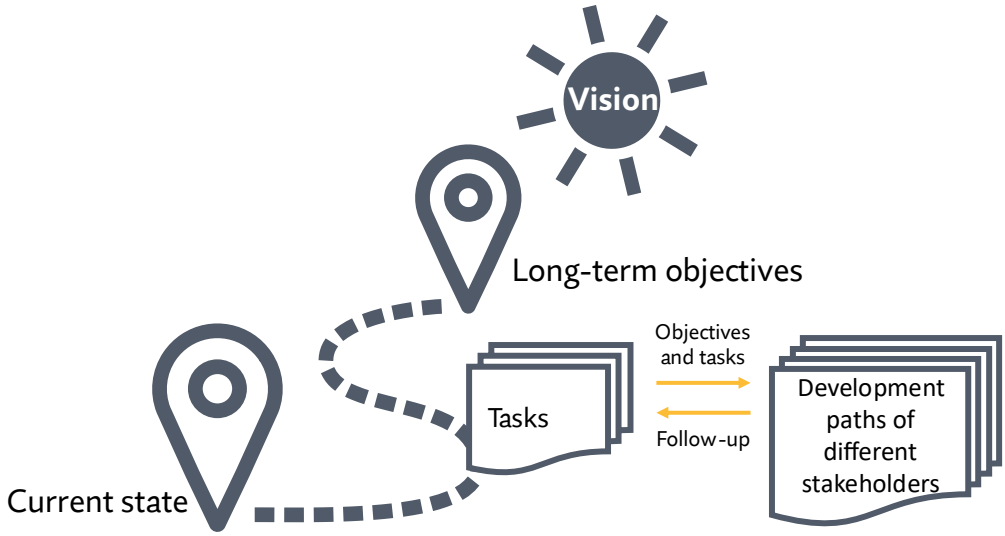

Figure 7. The structure of the GSFI development roadmap.

2.4.3. GSFI national maintenance model

GSGF Finland (GSFI) is intended to be a continuously evolving entity. The GSFI national maintenance model describes the management of the structure and content of GSFI, the management of the GSFI roadmap work, and the roles required for these tasks. The aim of the GSFI maintenance model is to ensure that the operating models and governance practices for the maintenance and development of GSFI are commonly defined, in use, and maintained. The maintenance model also aims to expand the utilisation of GSFI components throughout their lifecycle. Changes to GSFI and its' elements are processed and approved as described in the maintenance model.

However, the document is still a proposal and in this form is intended as a basis for discussion and building cooperation. Once the maintenance model has been completed and confirmed, it requires all parties to commit to the operations and roles described by the document. The proposal for GSFI maintenance model is described in the document GSGF Finland: National Maintenance Model (draft) (GSFI and Eurostat, 2024b).

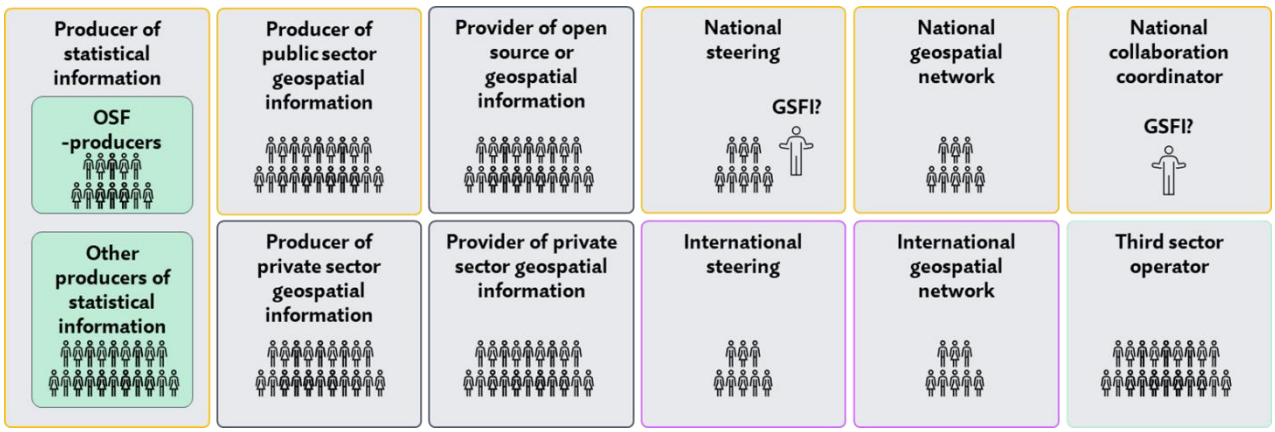

Figure 8. An illustration of the target state architectural actors and roles description. From the point of view of managing GSFI, it is essential to define the missing responsibilities for the roles of the national steering and collaboration coordinator of GSFI.

2.5. Tools

The common tools supporting GSFI-compliant activities are essential for creating an efficient and unified operational environment. This section covers the first tools already available, including a common glossary, architectural descriptions, a proposal for a small area division model, studies of confidentiality aspects and of data quality and value chains, and communication materials.

2.5.1. GSFI glossary

The Glossary of Geospatial Information in Statistics defines key concepts related to statistics and geospatial information. The aim of the glossary is to clarify the diverse use of terms and thereby support the collaboration of statistics and geospatial information experts, assist data producers and users in Finland, and enable data interoperability. The glossary is also hoped to be useful for communications officers, journalists, and translators.

The glossary contains information on 31 concepts related to geospatial information and statistics. The content of the concepts is described through definitions and supplementary information. Recommendations are provided for Finnish terms, and equivalents are given in English. The relationships between concepts are illustrated using concept system diagrams. The glossary was created in collaboration between Statistics Finland, the Finnish Terminology Centre, the National Land Survey of Finland, and the Finnish Environment Institute. The terminologist from the Finnish Terminology Centre was responsible for the terminological work. The glossary is available on the website of the Finnish Terminology Centre (GSFI and Eurostat, 2024c). Afterwards, the glossary was also translated into English, but it must be noted that the English version is not the result of the actual terminology work (GSFI and Eurostat, 2024d).

2.5.2. Enterprise architecture

The target state architecture descriptions provide structure and direction for development work. One of the primary benefits of architectural descriptions is the alignment it fosters between different parties. For example, with the help of the enterprise architecture, it is possible to define the desired future state in a structural way so that the national level descriptions can be analysed and combined with the organisations' own architecture descriptions and planning. This can influence the organisations' investments and initiatives to also support the GSFI objectives. This consistency also helps to prioritise projects that produce the most value and to ensure that resources are allocated efficiently.

The first target state architecture descriptions in the Finnish context produced for the GSFI entity are:

- Roles and Services (Figure 9): View of the spatial data infrastructure: This diagram presents the key roles for the spatial data infrastructure and the business services under the responsibility of these roles. Additionally, the diagram depicts the interactions between services through data flows.
- Layered View: This diagram refines the roles and services view of the spatial data infrastructure. It provides additional information about the business processes, information, and applications contained in the services in the target state. The layered view aims to identify the necessary elements (building blocks) of the spatial data infrastructure that allow location information to be linked as efficiently and consistently as possible to information and statistical production from the moment the data is created to its distribution. Roles are responsible for certain business services, the content of which is described through business processes and

business concepts. The bottom layer of the diagram consists of the application level, which can be supplemented in further work by identifying and describing general application services and data repositories.
The framework for the entire diagram is formed by different levels of interoperability: Legislation must enable interoperability at all levels. At the organisational level, it must be agreed in which roles the organisations operating in the ecosystem appear and on uniform operating methods. Semantic interoperability ensures that data contents are understood in a consistent manner, and technical interoperability ensures that data can be transferred in an interoperable way.

- Data Flow View: The purpose of the data flow view is to clarify the responsibilities between the roles in the spatial data infrastructure. It depicts what data can flow between roles and thus provides additional information to the layer diagram, which describes the infrastructure from the perspective of building blocks, not operations. In the data flow view, attention should also be paid to the process of validating location information. Validation is carried out using the service provided by the part of the spatial data infrastructure responsible for determining the location in the service of the source data producer. This way, validation is handled immediately when the data is created, and the organisations utilising the data can trust the quality of the received location information without needing to perform their own validation processes for this purpose.
- Actors and Roles View: This description shows how diverse the field of actors is in Finland. The roles presented in the description are based on the roles identified in the GSGF Europe work, to which actors present in Finland have been linked. It should be noted that no actors have yet been identified for the critical role of national coordinator, which is essential for the implementation of GSFI.
- Area Hierarchy View: A view that demonstrates the interrelationships between different areal divisions that can be used to present statistical data, as well as how different area divisions are related to each other. The area hierarchy view also shows how the base areas of the national small area division could enable a hierarchical system of area divisions in the future.

Architecture descriptions provide a good basis for discussions aimed at achieving a common understanding. As the group of participating actors expands and a better understanding is gained from these discussions, the descriptions will evolve and become more precise.

The architecture descriptions are available on the GSFI website (GSFI and Eurostat, 2024e). Additionally, GSGF Europe: GEOSTAT Information Service offers a set of reference architecture descriptions and reviews, such as linkages to GSBPM and GeoGSBPM (EFGS, 2024).

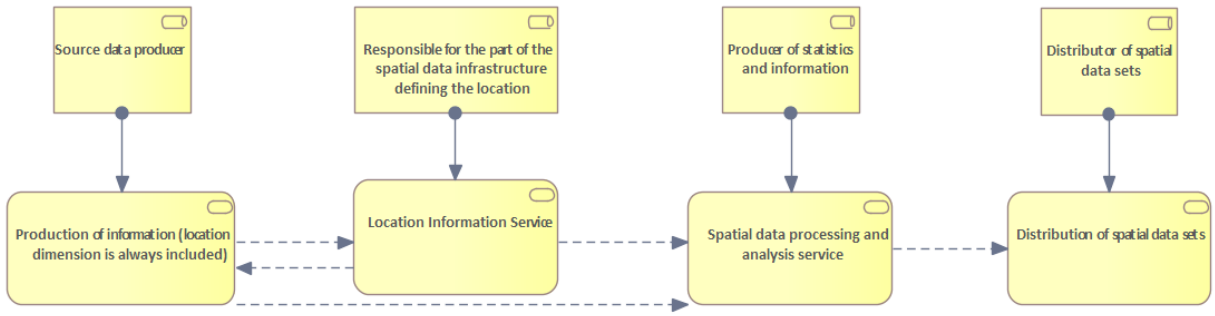

Figure 9. Roles and services of spatial data infrastructure in statistical production.

2.5.3. Proposal for a national small area division model

GSFI also provides a proposal for a new national small area division. A small area refers to a statistical unit that divides municipalities into smaller statistical areas. The development of small areas is linked to Principle 3 of the GSGF framework, 'Common Geographies for the Dissemination of Statistics.' Principle 3 ensures that common geographies are available for data dissemination. These support the management and use of data from different sources as geospatial information, as well as the integration, visualisation, analysis, and interpretation of data. In the GSFI roadmap, the proposal for a model of small area divisions is an important step in implementing action 3.1.1 'Define and describe common national geographies.'

The goal of the new division model is to provide a unified small area system and a corresponding classification system that is compatible with administrative boundaries and uniform across the country. The new small-area division model aims to complement municipalities' own operational divisions by providing a foundational dataset (base areas) for the flexible formation of different small-areas, also enabling the creation of nationally consistent area divisions.

Small areas and statistical grids are essential for area specific statistics and regional analysis. The variation of different variables occurs at different geographical scales, so the type of area unit needed depends on the phenomenon being examined. Currently, statistics smaller than municipalities are produced, for example, by municipal sub-areas, postal code areas, or statistical grids.

Each of these units offers unique advantages: municipal sub-areas provide detailed local insights, postal code areas offer a familiar framework, and statistical grids are very flexible, allowing for consideration of spatial extent of analysed topic as well as delimitation of the area under examination and cartographic representation of the phenomenon. Each of these also has limitations regarding national consistency, compatibility with administrative boundaries or data privacy.

The proposed new small area division model provides a nationally consistent and comparable dataset that enhances our understanding of regional dynamics while maintaining data privacy and compatibility with administrative boundaries. The model is based on an approach where a dataset of highly detailed base areas is first created. This dataset serves as the foundation for defining the actual statistical small areas. The base areas take into account the urban structure and geographical factors. Using the base areas, it is possible to create small areas based on various criteria. The definition of statistical small areas relies on population size criteria, ensuring data privacy when publishing statistical information.

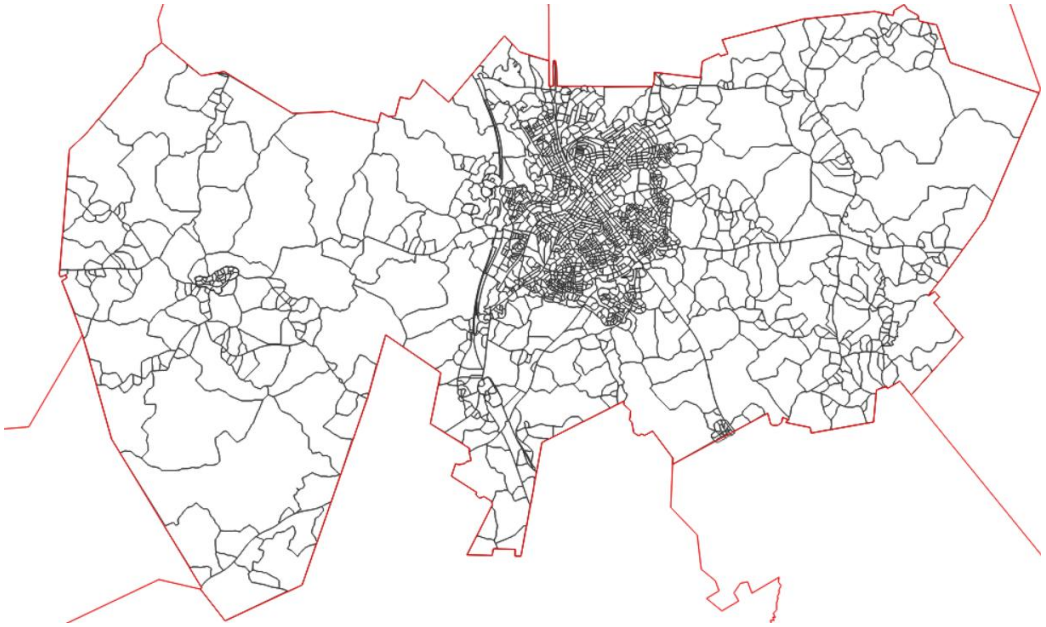

Figure 10. An example of base areas in Hyvinkää. Base areas are used as a foundation for statistical small areas. Figure illustrates the test version of base areas.

The needs assessment underlying the model development, the objectives set for the model, the testing of different approaches, and the proposal for the model of small area divisions are described in document 'Proposal for a new small area division model in Finland (GSFI and Eurostat, 2024f).

2.5.4. Data protection in the integration of statistical and geospatial information

When publishing statistical data, attention must always be paid to statistical confidentiality. The goal is to prevent the identification of individual statistical units, such as individuals or companies, and the disclosure of information about them from the published data. In Finland, Statistics Finland and other authorities combine statistical data with geospatial information to offer various services, such as the Grid Database and Monitoring System of Spatial Structure and Urban Form (YKR data).

The importance of data protection is emphasised when statistical data is combined with precise location information, making it possible to identify individual statistical units. The geographical location of a statistical unit can serve as an indirect identifier, increasing the risk of disclosure. To appropriately ensure data protection in the context of small area statistics, it is important to recognise these specific features and challenges related to data protection that precise location information poses:

- Risk related to the accuracy of geospatial information: The more precise the geospatial information, i.e., the smaller the identifiable geographical area the statistical data pertains to, the greater the risk of disclosing the statistical units behind the data. Statistical data should not be published if it pertains to only a single observation, i.e., a statistical unit, as the risk of disclosure for that unit is too high.
- Risk related to clusters of observations and autocorrelation: In spatially autocorrelated data, observations that are geographically close to each other receive similar values (positive autocorrelation) or, conversely, very different values (negative autocorrelation). Strong positive autocorrelation increases the risk of disclosure, as aggregates formed from the observations can be very close to the values of individual observations.

- Risk related to intersecting statistical areas: When data on the same topic is produced using different area classifications that are not nested, intersecting statistical areas may be created by overlaying the classifications. Different area classifications include e.g. administrative areas, grid data of various sizes, postal code areas, and electoral districts. If such an intersecting statistical area is very small, i.e., it includes only one or a few statistical units, publishing the data may lead to the disclosure of the identity and/or attributes of these statistical units.

Since geospatial information often increases the risk of disclosing statistical units, statistical disclosure control (SDC) methods should be applied to statistical data containing geospatial information to reduce the risk. SDC methods for aggregated data can be roughly divided into two categories: methods applied before aggregation (pre-tabular) and methods applied after aggregation (post-tabular). Pre-tabular methods are usually perturbative methods. They aim to alter the values of unit-level data so that aggregates formed from the perturbed data cannot be used to reliably infer anything about the original unit-level values. Post-tabular SDC methods are often restrictive methods. In other words, the methods are based on restricting the publication of data obtained as a result of aggregation if it is assessed to pose too high disclosure risk.

International recommendations emphasise increasing awareness of the specific features of geospatial information and considering them in publication and data protection practices. For example, the UN Expert Group on the Integration of Statistical and Geospatial Information (UN EG-ISGI) has made recommendations to address data protection and related challenges when combining statistical and geospatial data.

According to international recommendations, national legislation should consider the specific features related to data protection issues of statistical data containing geospatial information. For example, Statistics Finland's operations are strongly guided by the Statistics Act (280/2004), which obliges the protection of confidential information collected for statistical purposes. Geospatial information is not specifically mentioned in this law. However, in the government's proposal to amend the Statistics Act (HE 154/2012), geospatial information has already been highlighted.

In recent years, Statistics Finland has increasingly focused on the data protection of geospatial information and small area statistics. One reason for this is the current society's technological capabilities, which also bring challenges to secure data processing. The amount of open, high-quality, comprehensive, and combinable data is increasing, while geospatial applications are no longer only available to GIS (geographic information system) experts.

Currently, Statistics Finland uses restrictive SDC methods, such as (cell) suppression and aggregation of area data, to protect statistical data combined with geospatial information. In statistical production containing geospatial information, one-way distributions of variables are also preferred over cross-tabulation. Protecting statistical data combined with geospatial information is sometimes challenging. For example, decisions related to protection process may need to be made on a case-by-case basis, one assignment at a time. Although there are challenges associated with the use of restrictive methods, their use is well established at Statistics Finland, and transitioning to perturbative protection methods would be a significant change that should be carefully prepared, considering both the producers and users of statistical data.

Another example of the need for careful consideration of geospatial data protection relates to the openness of data in changing geopolitical situations. In addition to the integration of statistical and geospatial data, other open geospatial datasets and related risks have also received attention at the national level. In 2023, the Ministry of Finance established a working group to prepare a national risk

assessment and operational model for handling geospatial data (project number VM111:00/2023). A public version of the memorandum prepared by the working group has also been published, which mentions Statistics Finland's geospatial data as part of the key geospatial data resources of authorities (Ministry of Finance, 2024).

This topic has been discussed more broadly in a document on Data Protection in the Integration of Statistical and Geospatial Data (GSFI and Eurostat, 2024g).

2.5.5. Value chains and data quality

One challenge for the smooth integration and interoperability of statistical and geospatial information is ensuring seamless data flow between organisations and administrative sectors. Achieving interoperability requires cooperation both within and between organisations, as well as a shared understanding of development measures, their benefits, and agreed-upon implementations.

Typically, information introduced at one stage of a process does not meet the needs of the next stage and requires corrections and modifications. These points of value degradation can occur within an organisation's own processes but are more likely when information flows from one organisation to another (Figure 11). Value chains enable the description and understanding of changes in the value of information at different stages of the data flow. GSFI proposes utilising the national Data Quality Framework in value chain analysis, particularly focusing on stages where value degrades. The eleven quality criteria of the national data quality framework provide a standardised structure for assessing information quality at these points (Statistics Finland, 2022). The quality criteria are correctness, accuracy, consistency, currentness, completeness, traceability, understandability, compliance, portability, user rights, and punctuality (Figure 12).

Through value chain descriptions, it is possible to create a shared understanding of data flow and gain insights, especially at critical points in the process where information transitions from one actor to another or from one process stage to another. When discussing data quality, it is important to remember that the quality of information is determined by the needs of the data user. In the current situation, it is possible that the data producer is not aware of all the intended uses of the data or the specific needs of data users. Often, the data producer does not receive feedback on corrective actions taken by data users. The situational understanding formed through the data quality framework and value chain analysis can be shared with other stakeholders or process participants, leading to changes that eliminate value degradation or reduce flow impediments. As a result, data flow becomes more efficient, processes are streamlined, and data producers can genuinely believe that the quality of the data they provide is good. In the ideal state, the data used for statistical production is a seamless part of a broader information ecosystem. The necessary data is readily available, its content meets user needs, thus, the quality of the data is high, requiring no corrective actions.

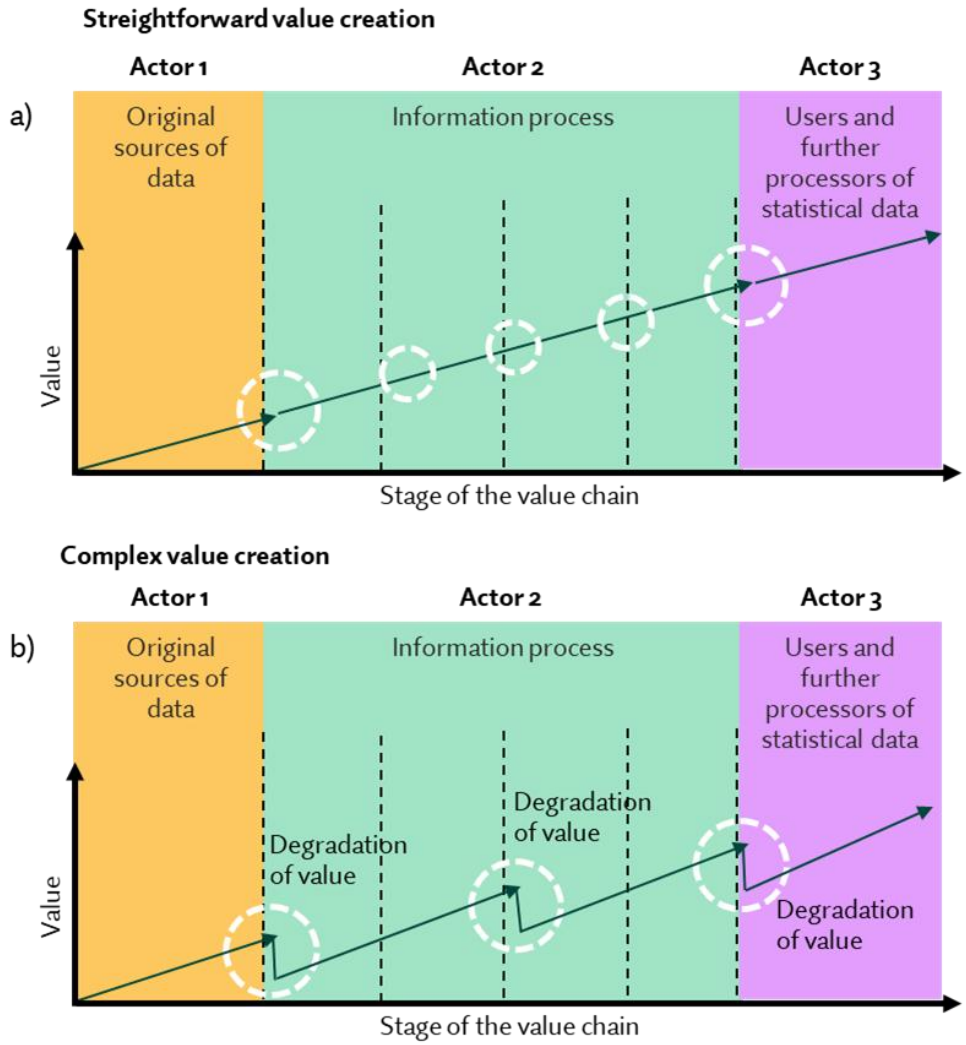

Figure 11. Straightforward and complex value creation. The value of information increases as it progresses through the information value chain. In a straightforward value chain, information advances without quality degradation, allowing it to be directly usable for the next actor or in the next process stage. In a complex value chain, the increase in the value of information is preceded by a decrease in value, which requires data processing. Through the editing processes, the value of information increases again.

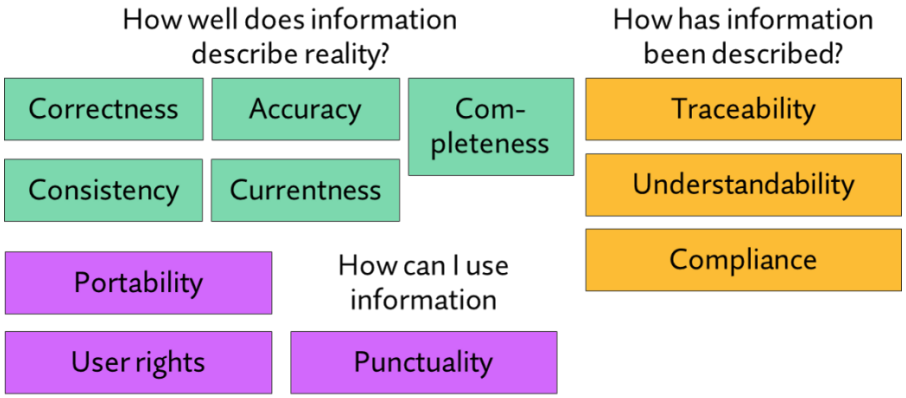

Figure 12. The national data quality framework: 11 quality criteria by category. When considering data quality, the key aspects tell the user what the data are about, and how well the data reflects reality. In terms of data usability, it is essential to know how the data can be used, and how well they are described.

This topic has been discussed more broadly in document on 'Practical approach in developing the integration of statistics and geospatial information through value chains and data quality' (GSFI and Eurostat, 2024h).

2.5.6. Communication

Effective communication is the cornerstone of any successful implementation. At GSFI, the aim is to ensure clear, consistent, and easily accessible communication for all interested parties. To facilitate this, communication platforms and tools have been developed, including:

- PowerPoint Presentations: Available in both Finnish and English, these presentations provide detailed overviews, ensuring that all participants have access to the same information and are also able to use the material in their own communication.
- Website: A website dedicated to GSFI, 'stat.fi/GSFI', serves as a central hub for all GSFI-related information and updates (in Finnish). In the future, there are plans to expand the current one page into a more comprehensive information service. It is designed to be user-friendly and easily navigable, ensuring that participants can quickly find the information they need.
- Teams Channel: A channel in Microsoft Teams is available to all members of the NISGIF network, fostering real-time collaboration and communication. This platform allows for instant messaging, file sharing, and virtual meetings, making it easier for participants to stay connected and commitment to the implementation of GSFI.

Visual Design and Templates: To maintain a consistent and professional appearance across all communications, a set of visual designs with an own logo and templates for Word and PowerPoint are developed. The available languages are Finnish, Swedish and English. These templates ensure that all documents and presentations adhere to the same standards of quality and branding.

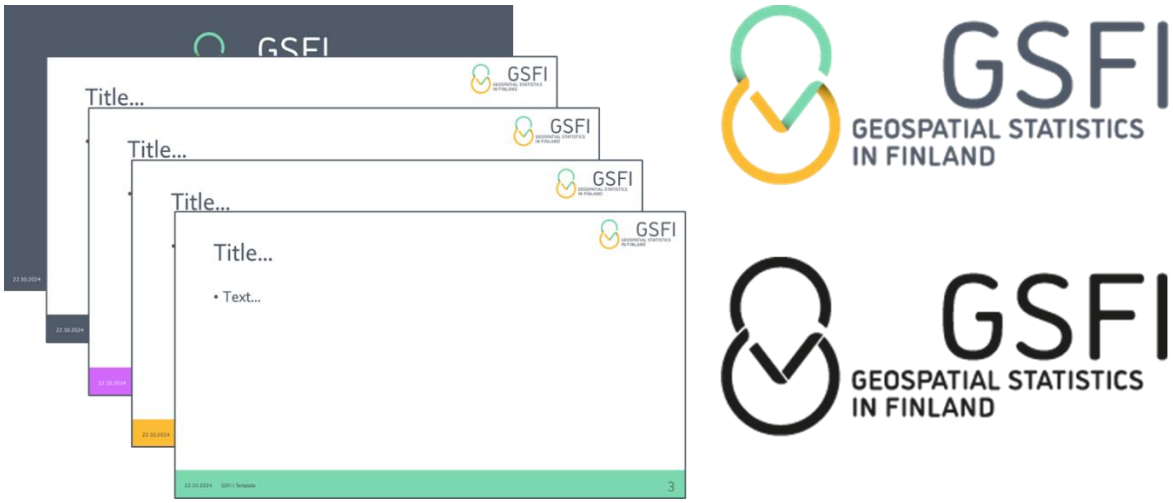

Figure 13. Geospatial statistics in Finland PowerPoint -templates and the logo.

2.6. Library

In addition to the tools, curated analyses, studies, and other information are collected to support the understanding of the operational field and the development of spatial data in Finland's statistical production. Currently available are a case study on the business data production process in Statistics Finland, and a present status review for the year 2023.

2.6.1. The Business Data Production Process, Statistics Finland – A Case Study

Practical case studies, such as this one at GSFI, provide valuable insight into design and streamlining of statistical processes in line with GSGF's statistics and geospatial interoperability solutions. The case study can also be used to demonstrate how geospatial information can be assimilated into statistical processes, improving efficiency and quality. In all cases, case studies bring important practical insight into planning, which otherwise would proceed on a relatively abstract level.

This case study examined the vision and an example of a solution for the integration of geospatial information in different statistical processes, utilising a centralized business data source and other data sources related to geospatial information. Furthermore, the case study contributed significantly to finding a practical design to integrate the data quality aspects in the geospatial statistical production process.

The work started by producing a description of the current state of the business data production process and continued by identifying development needs and describing a proposal for the target state. The identification of development needs and description of the target state has relied on the GSGF Europe data set and in particular the GSGF Europe – requirements and recommendations document. The integrity and comprehensiveness of the target state were tested with expert interviews. In addition, it was shown how value chains can be applied in defining process-related development needs and how quality aspects can be integrated into geospatial production processes by applying the value chains.

The results show how Statistics Finland has been able to enhance the integration of geospatial information into various statistical processes by utilising a centralised business data source and other linked data. The geospatial reference architecture and information system form a strong foundation for this integration, supporting the move towards centralised data repositories. In the proposal for the target state of the case study, it is suggested that the register's personnel, address, and location data would be based on the national geospatial infrastructure. This would eliminate the need for separate direct collections of location data and manual correction of address data, making the process both cost-effective and faster.

Currently, address data can be incomplete or erroneous, but in the target state, location data would be based on the building location data from the dedicated system called RYHTI. High-quality address data is critical for achieving the benefits of the target state, and here, establishment of a national address register has been identified as essential.

In the target state, the regional location data of organisations would be included in the national data ecosystem, and the ecosystem being developed in the Real Time Economy project would enable the sharing of basic information about legal units and their local units. The Tax Administration's Income Register could provide linkage data, allowing the determination of the location of organisational activities and personnel.

The case study process currently aligns with Statistics Finland's target state data process. With the proposed development suggestions, the inference process would become more efficient, particularly in terms of reducing manual editing and direct data collection needs. The estimated benefit for Statistics Finland would be an annual saving of 1–3 FTEs.

This topic has been discussed more broadly in a document 'Geospatial Information in the Business Data Production Process Finland – A Case Study' (GSFI and Eurostat, 2024i).

2.6.2. Present status review 2023

The construction of GSGF Finland (GSFI) began with an analysis of the present state in Finland in 2023. The aim was to describe the present situation and identify development needs essential for the successful and efficient implementation of GSFI. Observations were collected at the workshop of the Network on Integration of Statistics and Geospatial Information in Finland (NISGIF) on May 4, 2023, and through an expert survey. Participants evaluated and commented on Finland's present situation compared to the GSGF Europe: Requirements and Recommendations (EFGS and Eurostat, 2022), which outlines essential aspects for GSGF implementation in Europe. Successes and challenges were identified across all five principles.

The analysis showed that the national geospatial data infrastructure is robust due to the comprehensive implementation of the INSPIRE directive and information available through web services. However, the national geospatial infrastructure and geocoding required by GSGF principle 1 remain unstructured, leading to overlapping work in several organisations. Additionally, data governance responsibilities are not mutually agreed upon. Location reference information between registers is not managed nationally and information about data used as location references in different datasets is not readily available.

Individual organizations may already largely operate in accordance with the GSGF principle 2 in Finland. For example, at Statistics Finland, location data is part of the organisation's overall architecture. The unit-level statistical production model allows for flexible data management over time, and internal location data services could be smoothly replaced by national services. However, between organisations, solutions are organisation-specific, rely on varying principles, and are inefficient at the national level. Thus, Finland lacks a unified national production model for geospatial statistical data that would ensure regional and temporal consistency.

When comparing the present status to principle 3, it is noted that geospatial datasets based on municipality classifications are produced comprehensively. However, area divisions more detailed than the municipal level are problematic. Municipalities are responsible for municipal sub-regions, resulting in heterogeneous sub-municipal areas. Positive is that a new national network has been established to manage statistical classifications Additionally, GSFI is promoting the development and implementation of a new small area model, thereby advancing the definition of new small-area statistical classification for Finland. The use of statistical grids is strong and well-established.

Principle 4 aims for statistical and geospatial interoperability, a goal shared by the national network NISGIF. The GSFI project contributes to this work too, including the GSGF Finland architecture and developing a terminological glossary with approximately 30 terms. However, the network's work is voluntary, and mechanisms to agree on common models, standards, and practices and commit organizations to them are still lacking. Another challenge is unclear value chains of data. Organisations find it difficult to see their operations in relation to the whole, and inter-organisational workflows are minimal.

Looking at principle 5, it can be concluded that the implementation of the INSPIRE and PSI directives has created a strong data infrastructure and standards for dissemination, supporting the open distribution of geospatial datasets in Finland. However, INSPIRE metadata does not include metadata for statistical data, so it does not appear in searches. Changes in the security situation have also initiated the drafting of national data protection and security guidelines. Technical solutions for dynamically integrating statistics and geospatial data have been piloted, but progress towards actual implementation of services has been very slow. Meanwhile, organizations implement separate services

and may publish the same data with different specifications. Thus, there is no common national understanding of the entire set of data services, their information architecture, and the necessary metadata. A comprehensive data catalog for statistical and geospatial data is missing, resulting in a confusing supply of statistical geospatial data for users.

A summary of the review for each GSGF principle are provided in Appendix C. More broadly, this topic has been discussed in a document 'Present status and development needs of statistics and geospatial data integration in Finland in 2023' (GSFI and Eurostat, 2024j).

3. Collaboration Plan of the National Network NISGIF, draft

As the GSGF understands it, integration means pragmatic collaboration, knowledge exchange and shared services between the geospatial and statistical domains (GEOSTAT 4 and Eurostat, 2021). This understanding of strong links between statistical processes and processes of other national stakeholders led to the foundation of the National Network on Integration of Statistics and Geospatial Information in Finland (NISGIF) in 2021. Presently, the network brings together around 30 organisations and over 80 experts from state organisations, private companies, municipalities and universities. GSFI supports and enhances cooperation within the NISGIF and subsequently increases the commitment of the participating organisations to the common national goals.

The Network for Integration of Statistics and Geospatial Information in Finland (NISGIF) was established in 2021. The NISGIF network is a collaborative forum for solving everyday problems encountered in the context of integration of geospatial and statistical information. The network is a unique node of diverse and wide range of expertise, and it promotes cooperation and exchange of information. Therefore, it is essential to support an efficient and optimal operation of the network and to find the right fora within the network for working on different aspects arising in the implementation of the GSFI.

With all this in mind, the network was evaluated after three years of operation. A survey was sent to network members, and material and memos produced by the network in years 2021–2023 were analysed. In addition to this, the GSFI project worked in close cooperation with the network and the coordination group of the network and learned what works fluently and what does not. Below, the results are summarised from five perspectives, i.e. purpose of the network, structure of the network, operating principles, communication, and goals and monitoring. In addition, development proposals are presented to enhance collaboration and thereby also to support the implementation of GSFI.

This topic has been discussed more broadly in a document 'Development of the operation of the national Network for Integration of Statistics and Geospatial Information, proposal' (GSFI and Eurostat, 2024k).

3.1. The current state of the network's operation

The purpose of the network: The foundation of the network is to create a collaborative environment where the combined efforts of stakeholders achieve something unattainable by any single entity alone. This voluntary, equal, reciprocal, and self-directed partnership is deemed beneficial and optimal by all parties involved. What works well includes an agile approach, allowing flexibility and adaptation to rapidly changing demands of problem-solving, and low threshold working groups that address issues more as phenomena or topics rather than strictly defined tasks. Additionally, sharing knowledge through discussions, tips, skills, and experiences is crucial. However, several challenges remain, such as how to solve complex problems with numerous interdependencies comprehensively, or how the

network could share a practical, common understanding of where the network is aiming at. Identifying the challenges that the network seeks to address through collaboration is essential.

The structure of the network: The network is led by chairpersons and a coordination group with a defined set of tasks and an extensive list of responsibilities. This group meets regularly to guide and plan the network's activities. What works well includes an established, effective, and praised governance mechanism. The coordination group is open to active members and those willing to take responsibility, such as working group leaders. However, the composition of the coordination group rarely changes and has been dominated by three key agencies that led the founding process of the network. This suggests a potential need for renewal and role reassessment. The task list for the coordination group is long and requires clarification, rephrasing, concretisation, and better visibility of the working model to the network. Additionally, there are no direct resources allocated to the network.

The operating principles of the network: The network is open, voluntary, and based on self-registration. What works well includes a large membership base, with successful initiation and "recruitment" even during the pandemic. New members continuously join through word-of-mouth across various channels, driven by interest in the topics. Ease of joining without bureaucracy, attracts a diverse range of personalities and perspectives, fostering creative ideas. Innovations can emerge more easily as there are seldomly specific distinct expectations or pre-assignments from the organisations of network members. However, challenges include the lack of systematic utilisation of the members' skills with respect to network goal setting. Network members do not know each other well due to very few in-person meetings, and there is limited multilateral communication outside the network meetings. Expectations and benefits are perhaps not emphasised adequately, leading to loose and indistinct membership with no commitment of participants. Membership is characteristically personal rather than organisational, resulting in limited allocated resources for active participation. Furthermore, it is always not obvious whether the network fails or succeeds to meet participants' expectations.

The communication: The network has its own website with a visual identity, including a logo chosen by network vote. There is a slide deck on the website outlining the network's primary goals. The network's name is descriptive and conveys the network's core mission. Additionally, the Teams team, provided by Statistics Finland, available to all members of the network, has been set up for file management, discussions and meetings. Meeting minutes are recorded, and members' thoughts and feelings are gathered using electronic platforms, such as Mentimeter and Miro. Thus, basic visibility is well-managed, and the reputation of the network is good. However, there is no regular communication, such as newsletters. Other challenges include changes in electronic platforms, and they are also typically managed by only one organisation. There is also very little discussion outside the meetings, and collected materials are not curated. Better utilisation of accumulated knowledge and systematic presentation could increase efficiency and help new members.

The goals and monitoring: In the target state of the network, the roles of different stakeholders are jointly defined, allowing each organisation to focus more clearly on their area of expertise, and development is carried out through joint projects to secure resourcing. What has worked well, includes a systematic setting of goals through Enterprise Architecture (EA) analysis. This was done in an EA working group and the work was supported by the entire network. It can be realised, that the ambitious target state is consistent with the GSGF. However, concreteness of goals is still missing. There is a need to make the network's goals more visible, clarify their origins, how they are

implemented in network activities, how they are maintained and developed, and what is to be expected next. Clear progress tracking and metrics are also missing.

3.2. Development areas

Based on the analysis, three areas for development are proposed:

- A playbook for the network: A comprehensive guide that would outline procedures, best practices, and strategies for managing and operating the network effectively. The playbook would help to standardise the work and thereby e.g. to reduce the time needed to perform routine tasks. New members would also quickly get the basic idea of the network referring to the playbook. Clarification and concretisation of the task list of the coordination group could also help to recruit new members, as well as, to justify resource needs. The playbook can also support the organisation of more effective communication and activate members of the network by describing the rules for the goal-setting and tracking the progress.
- Responsibilities of the network in the national road map: Identifying the challenges that the network seeks to address through collaboration is essential. There is a need to make the network's goals more visible and link them to the national roadmap. This would provide the network a clear vision, help to plan and prioritise the work.
- Promotion and communication material to introduce the network: Readily available material, e.g. a pitch presentation facilitates the promotion and communication of the network in various contexts. Achieving the benefits of the playbook and the national roadmap is also a communication challenge. Communication material is needed in attracting and committing new members and building collaboration too.

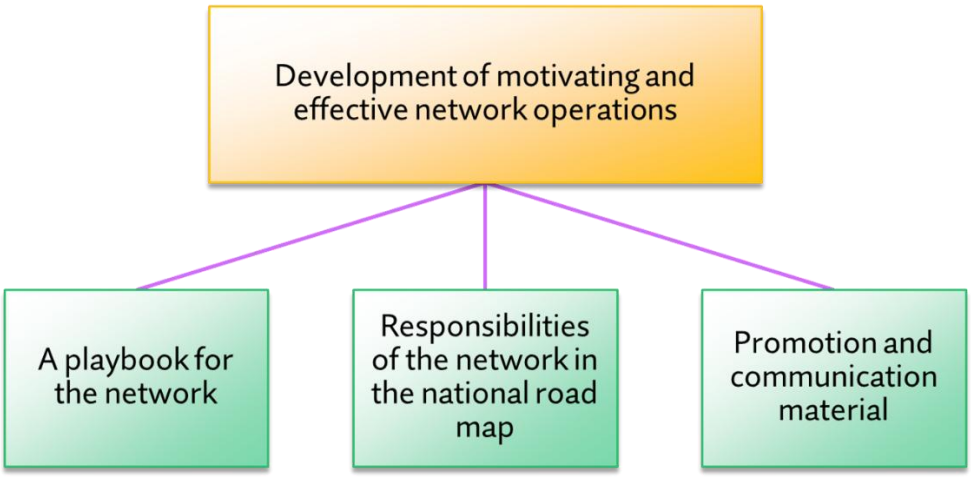

Figure 14. Based on the analysis, three areas for development are proposed.

In the playbook, it is important to clearly identify the different working forums and their roles so that each forum has its own place and predictable ways of working. This makes it easier to address topical issues in the network and help to choose the right forums to work. For the further discussion, the following forums have been outlined for the network.

General meetings of the network

- Presentation of various topics, plans, and results
- Regular dissemination of the GSFI and the present status in the development work

- Feedback and decisions using Mentimeter or similar tools

Goal-oriented working groups and smaller task forces

- In-depth work on a selected topic or phenomenon
- Can utilise other forums of the network

Coordination group

- Planning and coordinating network's activities, support working groups

Workshops, hackathons, and facilitated online sessions and meetings

- Identification of a specific problem or situation
- Brainstorming and planning solutions

Surveys and requests for comments to the entire network or targeted sub-groups

- Gathering, testing, and prioritising information
- Aiming for a comprehensive review of various skills and perspectives

Training webinars and online forums organized by experts in the network

- Cross-cutting or in-depth presentation of a specific theme or phenomenon

4. References

EFGS, 2024. GSGF Europe: GEOSTAT Information Service. A link to GSGF Europe: GEOSTAT Information Service web page in the online service of EFGS.

GEOSTAT 4 and Eurostat, 2021. GSGF Europe. EU. 945503 - 2019-FI-GEOSTAT4. A link to the GSGF Europe document.

GSFI and Eurostat, 2024a. GSFI development roadmap (draft). EU. 101112903 - 2022-FI-GEOS-GSFI.

GSFI and Eurostat, 2024b. GSGF Finland: National Maintenance Model (draft). EU. 101112903 - 2022-FI-GEOS-GSFI.

GSFI and Eurostat, 2024c. Glossary of Geospatial Information in Statistics. EU. 101112903 - 2022-FI-GEOS-GSFI. A link to the web page: Glossary of Geospatial Information in Statistics in the online service of Finnish Terminology Centre.

GSFI and Eurostat, 2024d. Glossary of Geospatial Information in Statistics. A translation of the Finnish glossary, not a result of a terminology work in English. EU. 101112903 - 2022-FI-GEOS-GSFI.

GSFI and Eurostat, 2024e. Geospatial statistics in Finland (GSFI). A web page. In Finnish. EU. 101112903 - 2022-FI-GEOS-GSFI. A link to the GSFI web page: stat.fi/GSFI.

GSFI and Eurostat, 2024f. Proposal for a new small area division model in Finland. EU. 101112903 - 2022-FI-GEOS-GSFI.

GSFI and Eurostat, 2024g. Data Protection in the Integration of Statistical and Geospatial Data. In Finnish. EU. 101112903 - 2022-FI-GEOS-GSFI.

GSFI and Eurostat, 2024h. Practical approach in developing the integration of statistics and geospatial information through value chains and data quality. In Finnish. EU. 101112903 - 2022-FI-GEOS-GSFI.

GSFI and Eurostat, 2024i. Geospatial Information in the Business Data Production Process Finland – A Case Study. In Finnish. EU. 101112903 - 2022-FI-GEOS-GSFI.

GSFI and Eurostat, 2024j. Present status and development needs of statistics and geospatial data integration in Finland in 2023. In Finnish. EU. 101112903 - 2022-FI-GEOS-GSFI.

GSFI and Eurostat, 2024k. GSFI and Eurostat, 2024. Development of the operation of the national Network for Integration of Statistics and Geospatial Information, proposal. In Finnish. EU. 101112903 - 2022-FI-GEOS-GSFI.

Tilastokeskus, 2022. Data quality framework. A link to the web page of Data quality framework in the online service of Statistics Finland.

UNSC and UN-GGIM, 2019. The Global Statistical Geospatial Framework. The GSGF. A link to the GSGF document (pdf).

Ministry of Finance, 2024, 2024. Memorandum of the National Geospatial Information Risk Assessment Working Group (public). In Finnish. VN/15876/2023. A link to the memorandum of the National Geospatial Information Risk Assessment Working Group (pdf). Referred 6.9.2024.

Ministry of Finance, 2020. Strategy for Public Governance Renewal. ISBN:978-952-367-690-9 A link to the Strategy for Public Governance Renewal on the web service of the Ministry of Finance. Referred 17.10.2024.

Appendix A. GSFI vision.

Table 1. GSFI vision.

Vision	Description
Finland has a national, widely used geospatial data infrastructure that provides location data for information and statistical production.	1. The key national data repositories are defined, and responsibilities assigned. They are the primary source of information in data processes. 2. Information needs are identified and described. When datasets or services are modified, the fulfilment of known needs is ensured, and changes are communicated.
In information and statistical production, location data is based on unit-level information, making integration with other data efficient.	1. Information and statistical production relies on the utilisation of data and services of the national data infrastructure. With the established information infrastructure, services, and agreed responsibilities, the information and services needed for the geospatial data-dependent phases of processes are comprehensively available, avoiding redundant efforts across organisations. 2. Information and statistical production is primarily based on unit-level datasets. These data units can be linked to a data unit that has location information. 3. The key national data units with location information include building and address data.
The national small area division model and statistical grids are the foundation of regional information.	1. The production of regional data and statistics more detailed than municipal divisions is based on the national small area system or statistical grids. 2. The maintenance and distribution responsibilities of the national small area system have been agreed upon. 3. New regional divisions more detailed than municipalities are always based on the utilisation of the small area division model.
Geospatial and regional data are discoverable and can be safely utilised.	1. Common data repositories, data identifiers, metadata, and secure data distribution technologies make information discoverable and accessible to users. 2. The integration of geospatial data and statistics is a natural part of national data services. Depending on the information need, data can be flexibly derived for different regional divisions. 3. Data protection is considered by data producers and services. There is guidance available on the data protection of regional information.
The significance of integrating statistics and geospatial information is recognised across administrative domains in society.	1. The broad scope of integrating statistics and geospatial data is understood. This is supported by materials describing the overall concept and highlighting best practices, which are made available to various stakeholders. 2. The cross-organisational model operates at different levels of activity; as expert-level collaboration, such as networks and joint projects, and as guiding activities across the administrative sectors of different ministries.

Appendix B. GSFI principles of implementation

Below the GSFI principles of implementation are listed in four tables. One table for each of the four key elements:

- Standards and good practices
- National Laws and Policy
- Technical infrastructure and capabilities
- Institutional collaboration

Table 2. Standards and Good Practices

Principle of implementation	Description
The implementation of the GSGF Finland is strengthened by developing statistical, data and geospatial standards and methods in collaboration and across sectoral boundaries.	The goal is to bring together perspectives and expertise from different sectors so that the standards and methods complement and support each other. The interoperability and boundaries of frameworks guiding different sectors are also challenged and developed.
The implementation of GSGF Finland relies on the description of activities.	The aim of the GSGF implementation in Finland is to achieve resource efficiency and interoperability through collaboration. Therefore, the description of activities and the sharing of information are the driving forces of the implementation. This support sharing of capabilities and execution of activities in a unified manner. The development is supported by a common reference architecture, the inclusion of geospatial aspects in the overall architecture of organizations, and the description of the methods, principles, and rules used.
Descriptions and identified best practices made during the implementation of GSGF Finland are communicated openly.	The accessibility of communication and information is increased by developing common sources of information, of which GSGF Finland entity is one. Important aspects include open information sharing, a well-thought-out structure, content curation, and descriptions that provide context and support usage.

Table 3. National laws and policy.

Principle of implementation	Description
The implementation of GSGF Finland relies on existing legislation and identifying areas for development.	The GSGF implementation in Finland is closely linked to legislation. Implementation requires an understanding of the opportunities and challenges that data legislation and its development present for the interoperability of statistics and geospatial information. Implementation may also need new support from legislation, for example, to ensure ownership and maintenance. Development is influenced together by communicating both the development needs and benefits.
The governance and funding solutions for the implementation of GSGF Finland extend across the administrative sectors of ministries.	Information processes cross administrative boundaries. To avoid sub-optimisation, various cross-organisational and cross-sectoral data value chains are considered in governance and funding. For example, effective financing solutions are sought for inter-organisational development. The benefits to be achieved are examined broadly, taking into account the entire public administration and society. It is continuously ensured that structures supporting cooperation and the capabilities corresponding to agreed responsibilities are found and developed. High-quality geospatial information is seen as one of the national key themes in the development of information interoperability.

Table 4. Technical infrastructure and capabilities.

Principle of implementation	Description
The implementation of GSGF Finland relies on continuous development, avoiding the accumulation of technical debt, and ensuring the interoperability of information systems.	Development is viewed ecosystemically, aiming for seamless linkage between different information systems, avoiding redundancies, and ensuring interoperability. The goal is to achieve advanced process automation and the utilisation of artificial intelligence. New technology is actively utilised and tested through various case studies. Progress and lessons learned are communicated within the information value chains.
The implementation of GSGF Finland takes into account the very different levels of geospatial expertise of the operators.	Geospatial expertise varies between operators. The implementation of solutions considers different levels of expertise and lowers the threshold for using geospatial infrastructure, with examples including supportive services, technical interfaces, guidance, collaboration, and peer learning, as well as providing training and open learning materials. (Geospatial expertise consists of knowledge in geospatial technology, geospatial methods and analysis, content management of geospatial datasets, information management, cartography, and data privacy in distribution.)
In the implementation of GSGF Finland, the capabilities of different actors are utilised for the benefit of all.	Geospatial capabilities are developed as a whole, taking into account both the different areas of capability and the possibilities for forming collaboration and skill pools. Geospatial expertise is integrated into other capabilities, understanding the spatial dimension of one's own substance and its management. (The three aspects of capability are i) operating models and processes, ii) personnel, skills, and resources, iii) data and systems.)

Table 5. Institutional collaboration.

Principle of implementation	Description
The implementation of GSGF Finland relies on operational models that support collaboration, active communication, and strengthened networks, as well as recognising roles of organisations and leading by example.	Effective collaboration in the implementation of GSGF Finland requires clear division of responsibilities, common operational models, principles, and places to work on issues together both vertically and horizontally. Attention is paid to workflows between organisations and common practices are agreed upon. Important is that organisations themselves set an example. When responsibilities and interrelationships are defined and described, it is easy to find the right contacts in each organisation for inquiries, communication, collaboration, etc., especially in inter-organisational and inter-sectoral development. Organisations spread knowledge about the benefits of collaboration and provide others with concrete opportunities for influence and participation. In ensuring planned and broad-based co-development, the important role of the integration network is recognised, and efforts are made to coordinate the network, work together, and disseminate results. Other expert networks are also strengthened and linked together.
In the implementation of GSGF Finland, public administration services and solutions are developed as a whole.	The implementation of GSGF Finland develops an interoperable decentralised system that broadly meets user needs. Solutions from one's own organisation are also shared for use by others. Collaboration between organisations avoids confusing and/or conflicting implementations from the user's perspective and reduces redundant work within organisations.

Appendix C. Present State in Finland in 2023, successes and challenges

Table 6. Successes and challenges at the present when compared to principle 1: Use of fundamental geospatial infrastructure and geocoding.

Successes	Challenges
• There is a wide variety of geospatial data available in Finland. • The implementation of the INSPIRE directive has improved the accessibility and descriptions of the data. • High-quality open location data is available, and licensing practices are clear. • The opening of geospatial data by the National Land Survey of Finland in 2011 was a significant step in improved collaboration. • New national location data registers, such as the RYHTI information system, are currently being developed.	• The absence of a national address register. • Location data in registers and the responsibilities of various actors are not structured and coordinated nationally. • Geocoding is carried out with various geospatial data and services, the usage of methods and purposes of which have not been defined nationally. • The management of the temporal dimension of location data is inadequate, e.g., the roles in archiving are unidentified. • Big data is not considered in the geospatial data infrastructure. • Data producers and users face extra and overlapping work, because of lower geospatial data quality. • International cooperation is extensive, but its results do not sufficiently disseminate within national networks and organizations.

Table 7. Principle 2, Geocoded unit record data in a data management environment.

Successes	Challenges
• Individual organisations can already operate largely in accordance with the GSGF. For example, at Statistics Finland, location data is part of the organisation's overall architecture, the unit-level statistical production model allows for flexible data management over time, and internal geospatial data services can be smoothly replaced by national services. • The experience and expertise accumulated within organizations should be utilised in developing the national model. • There is readiness to utilise national geospatial infrastructure services, such as the address register, instead of internal location data services.	• Finland lacks a unified national production model for geospatial data that ensures regional and temporal consistency. Solutions are organisation-specific and inefficient at the national level. • The relationship between statistical units and geospatial objects has not been commonly defined. • Data collection and geocoding rely on varying principles, location datasets, and services. • The level of geospatial expertise among data producers varies. • The processes for assessing and correcting the quality of location data are inadequate and fragmented.

Table 8. Principle 3, Common geographies for dissemination of statistics.

Successes	Challenges
• The situation of map datasets for municipality-based area divisions used in statistics is good, and the national division of resources is somewhat clear. • Statistics Finland maintains area classifications and is the geospatial data producer mandated by the INSPIRE directive. • The joint classification network of SVT actors started in autumn 2023. • The National Land Survey of Finland produces the base data for the map datasets used by Statistics Finland and publishes the INSPIRE administrative areas. • The use of statistical grids is strong and well-established.	• A common national model and infrastructure for area divisions are lacking. Solutions are organisation specific. For example, historical management is missing, and the situation of statistical sub-areas smaller than municipalities is varied and does not support the formation of national datasets. • Existing solutions are not well known; for instance, the differences between available map datasets can be unclear to users, and the availability of statistical grids as open data is not widely known. • The widely used postal code area division is not compatible with municipal boundaries or other statistical area divisions. • Data based on statistical grids is not directly compatible with administrative areas. • There is no geospatial data on voting districts in Finland. • Unique identifiers describing area divisions are missing. • Hexagons are not used.

Table 9. Principle 4, Statistical and geospatial interoperability.

Successes	Challenges
• The network for integrating statistics and geospatial data is working towards a common architecture and interoperability target state. • The RYHTI project promotes the usability of building information. • Statistical and geospatial glossaries have already been created, and the GSFI project continues to develop the integration glossary. • Open data and the INSPIRE directives promote the reuse of information.	• Conceptual work is slow relative to the rapidly changing operational environment. It is challenging to widely implement the work that has been done. • Mechanisms to agree on common models, standards, and practices, such as a common national architecture and the application of location data from different registers, are lacking. • There is no common plan for data publication. • There is room for improvement in the cooperation of public organisations and the shared use of interface services. Using the data often requires interpretation and versioning. • The lack of common national goals and division of resources weakens the interoperability of services. • An established operating model for co-development and comprehensive management has not yet been formed. For example, organisations find it difficult to see their operations in relation to the whole, and inter-organisational workflows are still minimal. • Good pilots have been implemented in Finland but getting them into production has been challenging. • Metadata for the data is partially described, and metadata models in different fields vary.

Table 10. Principle 5, Accessible and usable geospatially enabled statistics.

Successes	Challenges
• The opening of data in Finland has been strong since the early 2010. • The requirements of the INSPIRE directive led to the creation of a service infrastructure for data sharing. • Public organisations have opened data under the CC4.0 license. • Geospatial datasets have been cataloged and are discoverable through search services thanks to INSPIRE. • Drafting of national data protection and security guidelines has started.	• The changed security situation in Europe has led to a reassessment of open data licenses. • There are no common regulations or guidelines in Finland for the open sharing or licensing of geospatial data. Licensing practices vary. • Organisations implement separate services for customer needs and also publish the same data with different specifications, causing confusion from the user's perspective. • Services for dynamic data integration have been piloted but are not in production use. • There is no comprehensive data catalog for statistical and geospatial data. • INSPIRE mandates the cataloging and guidance of geospatial data, but metadata for statistical datasets is not included in the INSPIRE metadata model.