Merging statistics and geospatial information - introduction

Statistics Explained

This article forms part of Eurostat 's statistical report on the *Integration of statistical and geospatial information.* .

Location is a key attribute to virtually all official statistics: it provides the structure for collecting, processing, storing, analysing and aggregating data. Moreover, location is a concept that most people are comfortable with, as statistics for a specific place, region or area help people to understand the relevance of particular indicators.

Integration of statistical and geospatial information

Traditionally, geospatial information and statistics were attributed to different organisational entities in each country with little cooperation. However, the association of geography and statistics generates information far beyond the simple representation of data on a map. Linking geo-referenced and numerical statistics in spatial analysis reveals relationships and phenomena which are difficult to discover by analysing statistical databases alone. Technological advances and new policy demands have shown that both fields can beneficially be combined. This development has created organisational challenges for the European statistical system (ESS) , driving increased levels of cooperation between statistical authorities across the European Union (EU) and other organisations, such as national mapping and cadastral agencies, providers of big data or national space agencies.

What is a geographical information system (GIS)?

A geographic information system (GIS) is a tool for the management, analysis, presentation and dissemination of geo-referenced data, in other words, data associated to their geographic location. This is evidently the case for topographic information about roads, rivers or administrative boundaries which have been traditionally represented on maps. However, a wide range of additional data sources can also be geo-referenced. Indeed, all statistics inherently have a geographical dimension, be it data covering the whole of an EU Member State, a region, a smaller administrative unit, or indeed an enterprise or a household.

A data revolution has resulted in the ever-increasing availability of statistical and geospatial data. This pattern of development is linked to the growing volume of data that is generated by the internet of things. At the same time as the volume of geospatial data has been increasing exponentially, the ESS has undergone a comprehensive process of reform covering most aspects of its statistical production. These reforms have been driven, in part, by new demands from policymakers to support evidence-based decision-making through better descriptions of societies, economies and the environment within the context of, for example, globalisation, demographic challenges or environmental threats.

Policymaking has increasingly moved across the confines of national borders: sustainable development goal (SDG) indicators are an example of a current cross-border policy. At the same time, European funding for regional and cohesion policy has focused attention on specific territorial characteristics, for example targeting economic, environmental and social problems in cities and/or rural areas. Another change has been the increased level of demand for territorial disaggregation within official statistics: for example, citizens are often most affected by decisions which influence their immediate neighbourhood; this has resulted in demand for information at a very precise level of detail in order to analyse and illustrate the impact of various programmes and policies. As a result, policymakers and analysts are looking for detailed information across a broad range of spatial dimensions, such as cities and/or rural areas, local administrative units and even more detailed at the 1 km^2 grid cells like in the 2021 census results.

International background

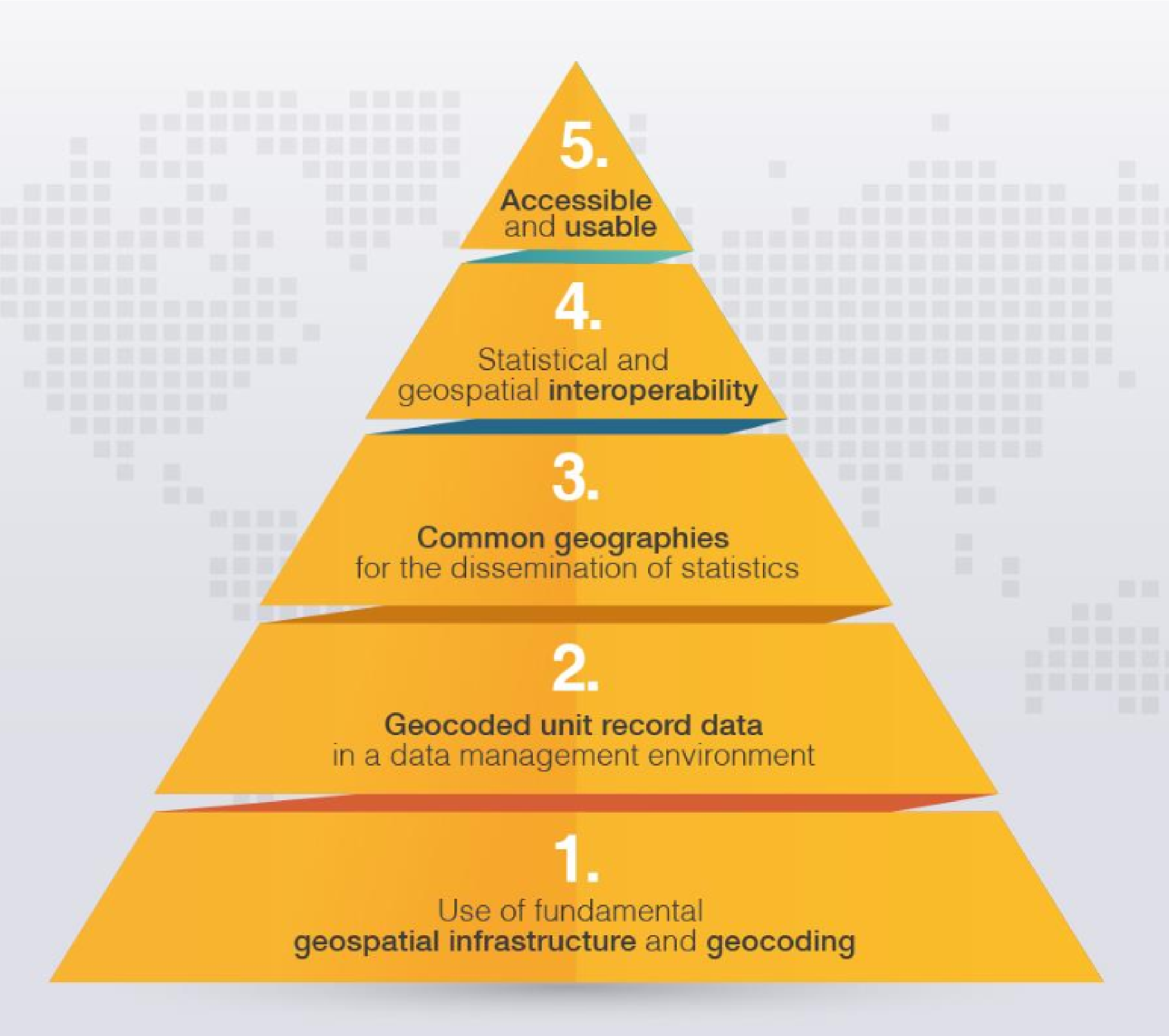
Globally, the lead on geospatial information is taken by the United Nations Committee of Experts on Global Geospatial Information Management (UN-GGIM) who acknowledged the 'critical importance of integrating geospatial information with statistics and socio-economic data'. Their work is based on the development of a global statistical geospatial framework (GSGF) designed to provide an interoperable method for geospatially coding and managing geospatial statistics and information, by connecting statistics that describe socioeconomic and environmental attributes to information that describes our physical man-made and natural environment. Eurostat is an observer of the UN-GGIM: Europe Executive Committee.

Additionally, the United Nations Integrated Geospatial Information Framework (UN-IGIF) provides a basis and guide for developing, integrating, strengthening and maximising geospatial information management and related resources in all countries. It will assist countries in bridging the geospatial digital divide, securing socio-economic prosperity, and leaving no one behind.

Within Europe, the implementation of a strategy for the integration of statistical and geospatial information follows global guidelines and has been organised, to a large degree, under the auspices of GEOSTAT projects and annual conferences of the European Forum for Geography and Statistics (EFGS) , both of which provide methodological guidance and are funded by Eurostat. The first GEOSTAT project was launched at the beginning of 2010 to promote grid-based statistics and was followed by three subsequent projects that focused on: a model for a point-based geocoding infrastructure; fostering better integration of geospatial information and statistics; developing a European version of the global statistical geospatial framework (GSGF Europe).

The global statistical geospatial framework (GSGF)

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



Source: United Nations Department of Economic and Social Affairs, United Nations Statistics Division, United Nations Committee of Experts on Global Geospatial Information Management and United Nations Statistical Commission (2019)

Global statistical geospatial framework pyramid

- Principle 1 specifies the adoption of a common and consistent approach to place each statistical unit of a dataset in time and space, using fundamental geospatial infrastructure.
- Principle 2 supports the process of linking or storing high-precision geographic references (in other words, geocodes – coordinates, small geographic area codes, or linked-data identifiers) to each microdata / statistical unit record. This is often referred to as geospatially enabling data, and must occur within a secure, standards-based data management environment.
- Principle 3 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 4 defines the preconditions for statistical and geospatial data to work as a data ecosystem, in which those involved interact with each other to exchange, produce and consume data. Interoperability between statistical and geospatial data and metadata standards is needed to overcome structural, semantic and syntactic barriers between data and metadata from different communities and providers.
- Principle 5 highlights the need for data custodians to make geospatially enabled statistics accessible and usable according to agreed standards and good practices, so that data users can discover, access, integrate,

analyse and visualise this information seamlessly for geographies of interest.

Source: The Global Statistical Geospatial Framework; United Nations Department of Economic and Social Affairs, United Nations Statistics Division, United Nations Committee of Experts on Global Geospatial Information Management and United Nations Statistical Commission (2019)

The ESS has responded to these challenges by acknowledging demand for a higher level of geographical detail in its statistics: identifying geospatial information as a valuable data source; recognising its potential for integrating data from multiple sources (administrative, statistical and/or big data); and acknowledging that it represents a considerable opportunity to create more relevant information and respond better to user needs. By geocoding various types of data, statisticians aim to link different datasets through the use of location as a neutral concept, thereby joining data disparate sources together. By doing so, the geocoding of data should help policymakers and analysts to answer the 'where?' in addition to the 'what?' and the 'when?' which have traditionally been the focus of official statistics.

Within national statistical authorities, competences for geospatial information often remains concentrated in specialised teams. They may have considerable experience in using geographical information systems (GIS), with the preparation of statistical maps an established practice. Some EU Member States also use maps to show statistics at a regional level or for smaller-scale territorial typologies. Hence, GIS is by no means a new technology for some members of the ESS. However, given the growing importance of geospatial information, the rapid expansion of data visualisation technologies and the emergence of entirely new groups of users, it would appear all the more important that the expert knowledge embedded in most national statistical authorities is disseminated as widely as possible to colleagues in the same Member State, as well as to other Member States. Yet, at the time of writing, the work conducted by a relatively large number of statisticians has yet to be impacted by developments in GIS, with many having few or no specific competences in this area. The implementation of geocoded unit record data and the use of fundamental geospatial infrastructure is still to be completed across many domains.

What is GISCO?

The geographical information system of the (European) Commission – GISCO – contains essential datasets such as topographic data and political boundaries. It provides corporate level data and services for a variety of users from the ESS and the EU institutions. Examples include data on administrative and statistical areas; population distribution, background maps or geocoding services. These are areas of common interest for multiple stakeholders and are regularly used by a range of European Commission services.

More information is available at: `https://ec.europa.eu/eurostat/web/gisco`

The focus of geospatial statistics is principally on spatial objects, in other words, real-world phenomena that have spatial extent or position. A spatial object has various characteristics (such as its shape, name or boundary). Spatial objects can be material (such as monuments, buildings or bridges) or non-material (such as administrative boundaries; boundaries of cadastral parcels of land; routes within a transport network). To code data at such a precise level, an address is often required: this provides a geographic data item that is used upstream of statistical processes, for example, during the collection of information either by post or in the field (by an interviewer). In both cases, an address provides direct access to the required respondent or object (by referring to a precise location with geographical coordinates that eliminates any ambiguity) – thereby establishing a geocoded unit record data entry.

An effective geo-referenced statistical information infrastructure should be consistent and interoperable with spatial data infrastructures developed following the INSPIRE Directive (see box below). Indeed, more accurate and better exploitable geospatial statistics may generate added value in a range of areas:

- statistical authorities may reduce the cost of data collection, for example, by using GIS techniques and location-enabled devices to plan surveys better;
- statistical collection could be made more efficient by exploiting geo-referenced data and/or making use of it when there were changes for reporting requirements;
- dissemination could be improved, by aggregating the same data to produce statistics for, among others, grids, functional areas, regions or river basin districts.

What is INSPIRE?

The INSPIRE Directive (2007/2/EC) entered into force in May 2007, establishing an infrastructure for spatial information in Europe (INSPIRE) to support Community environmental policies, and policies or activities which may have an impact on the environment. Its goal is to make geographic information held by public administrations more accessible through a geoportal (http://inspire-geoportal.ec.europa.eu/) that is accessible to everybody. To do so, data and metadata across 34 spatial data themes from regional, national and international sources are harmonised using an agreed set of standards that make it possible to share, combine and aggregate spatial information.

More information is available at: https://inspire.ec.europa.eu/

Investment in geospatial statistics also has the potential to improve greatly the synergies between national statistical authorities, mapping agencies and other authorities dealing in geographic and cadastral information: promoting the exchange and integration of statistics; avoiding duplication in data collection; and facilitating European, national and subnational reporting obligations. In order to be as effective and efficient as possible, the joining together of statistics and geospatial information needs to be focussed on harmonisation at the start of the process. In this way, geo-referenced statistics at geocoded unit record data (for example, microdata) level can later be fully exploited, allowing them to be used in more flexible ways and for a broader range of analyses. For example, a single set of geocoded unit record data may be used for providing statistics across a range of different territorial typologies, while the same data may also be used to analyse issues which are not necessarily known at the moment a survey was designed/carried out.

Such changes in statistical production have led to a number of established statistical conventions and rules being re-examined, the most prominent of which is the question of confidentiality. Official statistics have traditionally put considerable efforts into making the identification of statistical entities impossible or at least very difficult; this has mainly been done by eliminating confidential data from published results or aggregating data to such a level that prevents disclosure of the sensitive records. Geospatial information has a number of specific considerations and requires different types of safeguards to protect confidential information; the increased use of registries, disaggregation techniques and methods for modelling small areas are some of the most relevant areas of debate. Otherwise, the use of richer, more specific and detailed information also raises issues around business models and commercially relevant data, for example, those associated with common licencing schemes.

Purpose of this publication

Since 2009, Eurostat and the ESS have stepped up efforts to include detailed location as an important parameter in various environmental, social and economic statistics. This goal was designed to enhance the information capacity of statistical data, mainly for planning, programming and spatial analysis, without increasing the cost of creating the data; the main driver for this initiative was the 2011 census round.

Eurostat subsequently sought to promote integrated information systems that combine statistical and geospatial information for policymakers, researchers, spatial planners, as well as a range of other users, sharing this knowledge with the wider community, providing an overview of how geographical information systems have been implemented, increasing knowledge via the European statistical training programme (ESTP) and identifying issues for further guidance and future developments.

Another element was the launch of a call for proposals in 2012 under the heading of*Merging statistics and geospatial information* . This was designed to provide grants for facilitating work on the coordination of statistics and geospatial information. It was intended to cover a wider range of topics, including:

- improving the integration of geographic information and geo-referencing in the statistical production process;
- illustrating how linking geographical and statistical information provides additional value and creates new information;
- designing innovative web applications to show the spatial distribution of statistics.

The call was also intended to increase cooperation between national statistical authorities, in the sense that tools or processes designed or developed by one EU Member State might be offered to others for reuse and/or inspiration. Furthermore, national statistical authorities were explicitly encouraged to propose projects together with organisations responsible for geospatial information, in particular, national mapping and cadastral authorities (NMCA) to promote greater cooperation and a cross-fertilisation of information. Over time, the focus of the grants has evolved.

This publication, *Integration of statistical and geospatial information — experiences and observations from the national statistical authorities* , presents details for each of the projects provided a grant during the first nine years, showcasing a broad range of applications that may be developed using geospatial information.

For the 2012 exercise, Eurostat received 11 proposals and selected eight of these for grants, namely those from Greece, Hungary, Malta, the Netherlands, Poland, Slovenia, Slovakia and the United Kingdom. Some of the projects were cross-cutting and ranged from data collection to web dissemination, while others were focused on a specific aspect of the business process in an individual statistical authority. All projects were thought to have enhanced the GIS expertise of national statistical authorities and made substantial progress in giving increased visibility to GIS. However, there were concerns raised as to the potential transferability of projects between national statistical authorities, while no projects were carried out in unison with national mapping and cadastral authorities. In response, Eurostat set-up a collaborative platform to exchange information and promote reusing results from other countries. Furthermore, it was agreed that future calls should promote projects that sought to: bring location into the mainstream of statistical production (by developing geocoded data warehouses); help national statistical authorities prepare for the 2021 census exercise (through the implementation of geocoded data, covering buildings, addresses, citizens, businesses, workplaces and farm holdings, thus creating a point-based framework for statistical microdata).

For the 2013 exercise, there were seven projects selected by Eurostat for grants, namely those from Bulgaria, Germany, Croatia, Italy, Austria, Slovenia and Finland. One of these continued work that was started under the 2012 grant, whereas three others were centred on extending national capabilities for merging statistics and geospatial information. There were also three more specific projects that were focused on: manipulating information collected from migrant arrivals (collected when they applied for a residence permit), so that their place of birth could be geocoded, allowing a set of 20 maps to be produced for non-EU countries, showing the precise origin of migrant arrivals for the years 2012-2015. Another project allowed a set of commuter statistics to be developed, providing information on the average distance commuters travelled to work or to their studies. The final project was more closely linked to information technology, namely, developing an open-source web application for the spatial analysis of statistical data.

For the 2014 exercise, there were seven projects selected for grants, namely those from Estonia, France, Croatia, Hungary, Poland, Portugal and Norway. These covered a broad range of issues including: linking statistical registers to address systems; producing multi-modal spatial transport data for urban centres; assessing how changes in population and land use may impact on the quality of life; or establishing a point based business register.

For the 2015 exercise, there were eight projects selected for grants, namely those from France (this project was extended until 2018), Croatia, Latvia, the Netherlands, Austria, Poland, Slovenia and Finland. Note that the grant provided to the French national statistical authority concerned the preparation of a methodological handbook. As such, it did not specifically cover a practical application for merging statistics and geospatial information and for this reason has not been included in the main body of this publication. Nevertheless, the handbook that was produced provides a valuable tool that may be used to promote and share results, encouraging a greater take-up and application of spatial statistics in statistical production processes. The core of the handbook focuses on describing geocoded data, measuring the importance of spatial effects, describing practical methods for taking into account spatial interactions and providing details on some more advanced issues and latest developments (spatial panel data models, network analysis, spatial econometrics, small area methods). The *Handbook of Spatial Analysis* is available in both French and English.

For the 2016 exercise, five projects were selected by Eurostat, namely those from Greece, Italy, the Netherlands, Slovakia and Finland. For Italy, Slovakia and Finland, these were centred on extending national capabilities for merging statistics and geospatial information. The Greek project was focused on preparation for data collection for the 2021 population census, while the Dutch project analysed developments in paving (sealing) of land.

For the 2017 exercise, eight projects were selected, namely those from Denmark, Estonia, Croatia, the Netherlands, Austria, Poland, Finland and Norway. For Denmark, Croatia and Finland, these were centred on extending national capabilities for merging statistics and geospatial information. The Polish project was focused on evaluating the impact of changes in the boundaries of administrative units. The Estonian project was focused on developing indicators to improve local decision making. The Dutch, Austrian and Norwegian projects were focused on geospatial indicators for particular domains, energy and sustainability in the Netherlands, sustainable development indicators in Austria and housing and health in Norway.

For the 2018 exercise, there were five projects selected for grants, namely those from Poland, Finland, Sweden, Iceland and Norway. The Polish study was focused on developing a statistical disclosure strategy for geospatial data. The other studies were focused on geospatial indicators for particular domains: crop yields in Finland, economic activities (for example by businesses) in Sweden, register-based census statistics in Iceland, and land use/cover, crop identification and planning of urban developments and built-up areas in Norway.

No grants were issued in 2019. Attention was focused on a reflection/assessment of the grants issued between 2012 and 2015.

For the 2020 exercise, there were nine projects selected for grants, namely those from Bulgaria, Germany, Hungary, Malta, the Netherlands, Austria, Slovenia, Finland and Sweden. For Bulgaria, Germany, Hungary, Malta, the Netherlands and Finland, these were centred on extending national capabilities for the integration of statistical and geospatial information. Two studies were focused on geospatial indicators for particular domains: land cover statistics in Austria and agricultural land in Slovenia. The Swedish project was focused on identifying ways to improve geospatial quality within the Swedish business register.

Eurostat's role: the integration of statistical and geospatial information within the European statistical system

While geography has a long history in the statistical production process, globally, across the EU, nationally and regionally, there is a growing need to redesign official statistics by promoting – among other aspects – high-quality data with increasingly fine geographical detail to support international, national and regional policymaking. At the same time, statistical users in today's world expect to have georeferenced data, with more detailed (granular) and flexible referencing systems. This has led statistical authorities to consider location information and spatial analyses as essential elements for the development of modern statistical systems.

Through different calls for proposals, Eurostat has funded development work around the integration of statistical and geospatial information, helping statistical authorities in the EU Member States and EFTA countries to develop data infrastructures and supporting frameworks. Since 2010, Eurostat has supported four separate GEOSTAT projects with the goal of building a methodological foundation for the integration of statistical and geospatial information, while promoting the production of geospatial statistics in the European statistical system (ESS). These projects built on successive outcomes, expanding in both scope and ambition. They moved from an initial focus on practical use cases, applied in a harmonised and consistent way, to more general guidelines that encompass information, processes, technological and organisational aspects, which are expected to be drivers for the future of statistical and geospatial data integration in the ESS.

The funding provided by Eurostat for the grants under the topic 'Integration of statistical and geospatial information' has opened up opportunities to reuse and adapt global models and frameworks within the European context, identifying synergies, mutual gaps and opportunities. It has promoted cooperation between different stakeholders, driving the production of high-quality statistical products that more closely meet user needs. By doing so, Eurostat has helped to develop new geospatial information capacity that can be used to monitor strategic objectives and policy instruments globally, across the EU, nationally and more locally. Examples include information for the United Nations' sustainable development goals (SDGs) and the European Green Deal , or initiatives to build A Europe fit for the digital age or an Economy that works for people .

Eurostat funding has enhanced geospatial data availability and cooperation between statistical authorities and national mapping, cadastral and land registration authorities. At the same time, it has promoted the use of Earth observation for developing, complementing or enhancing statistical production processes. Since 2020 additional support has been provided by DG REGIO to support the implementation of the integration of statistical and geospatial information.

Ambition

When assessing the potential to award funding, Eurostat sought to measure the potential of each project by objectively referring to a list of criteria, including the:

- novelty, complexity and sustainability of linking/integrating geospatial information or Earth observation sources and statistical data;
- innovativeness of the approach adopted;
- likely quality of the derived information;
- potential and/or effective solutions delivered to the national statistical infrastructures;
- extent to which methodology and/or software solutions could be reused by others;
- provision for free and open access to data, algorithms, and software developments;
- quality of documentation, reports and studies, to allow for uptake by other organisations (in the same country or another EU Member State).

Nationally, the integration of geospatial, Earth observation and statistical data varies significantly across EU Member States. Eurostat's funding programme has sought to 'level the playing field', by providing support to those statistical authorities that need to develop their basic statistical infrastructures. However, funding has also been provided at the opposite end of the spectrum to those statistical authorities that are among the leading exponents of geospatial statistics in the EU, to encourage them to embark on innovative development work.

Eurostat's strategy of providing funding and supporting the integration of statistical and geospatial information was reinforced during a 2021 conference on Earth observation for official statistics, when the statistical community acknowledged the need for such efforts to make use of Earth observation data in official statistics. In line with the conclusions from an ESS meeting of 9 February 2012 and the adoption of the 2021 Warsaw memorandum on Earth observation for statistics , Eurostat continues to fund a wide range of projects that seek to link statistical and geospatial information, including Earth observation.

The funding provided by Eurostat has sought to provide support for the integration of geospatial and statistical information within national production processes, considering already existing methodologies and experiences from previous projects and various internationally recognised frameworks (IGIF, GSGF and its European derivative GSGF: Europe). It has provided financial support to the development of Earth observation data, tools and techniques for developing, complementing or enhancing statistical production processes. The work carried out has also allowed processes to be developed that establish regular updates of geospatial statistics that are used in evidence-based policymaking for local, subnational, national, EU and global development priorities/agendas. A final report was produced for each project that received funding, facilitating the exchange of information and best practice, allowing reuse of the processes and knowledge, while providing a support network for the development of geospatial statistics between EU Member States organisations.

European added value

The supported projects advanced the development of geospatial statistics in the EU, promoting standardisation for the collection, processing, analysing and dissemination of geospatial data. These works have enhanced the availability, accuracy and usability of geospatial data and statistics for a wide range of stakeholders. They encouraged statistical production to be based on common sources of geospatial information (e.g. a geocoding services), thereby leading to a more efficient use of resources and improved integration of data produced by different public organisations. Indeed, Eurostat promotes initiatives that seek to avoid the duplication of work, promoting data sharing, the production of data that is fit for purpose in harmonised formats, with the necessary spatial and temporal resolutions.

Impact

By adhering to established standards and best practices, Eurostat has promoted the accuracy and credibility of geospatial data in the EU, while aiming to ensure the quality and reliability of geospatial statistics. At the same time, Eurostat has supported capacity building efforts by providing training and technical assistance to national statistical authorities and other stakeholders through workshops, seminars and knowledge-sharing initiatives. Eurostat has fostered innovation and research in the field of geospatial statistics, collaborating with statistical authorities, academic institutions and research organisations to develop new methodologies, tools and applications.

Conclusion

Eurostat's role in advancing the integration of statistical and geospatial information within the EU has been instrumental. As a result, the funding provided by Eurostat has led to the delivery of better geospatial statistics across Europe, with improved timeliness and quality, bringing all contributors to the same level. Efficiency gains through investment in new processes have resulted in faster and more accurate data (with better error-checking). These processes have also reduced the burden on some data providers and new forms of analyses have become available by connecting disparate data sources to geospatial identifiers.

Glossary

Address is the specific location of a property, usually based on address identifiers such as a road name, house number and/or postal code.

Application programming interface (API) is a type of software interface used for computer programmes to communicate with each other.

Common geographies are consistent and accessible administrative units.

Continuous data describes data where values for the variable of interest may be observed at any point across the territory studied. Data are generated on a continuous basis, but they are measured only at a discrete number of points (for example, the chemical composition of the soil, water or air quality when analysing land use or land cover).

Geocoded unit record data are data that have been linked – through a set of geocodes – to a location. The geocode information can be used to combine unit record data, for example for analysis or reporting.

Geocoding is the process of transforming a description of location (such as an address or the name of a place) to a location on the Earth's surface. Geocoding is the process of linking unreferenced location information, often in the form of a text string (an address) to a geocode. The conditions for geocoding include a high-quality physical address, property or building identifier, or other location descriptor, in order to assign accurate coordinates and/or a small geographic area to each statistical unit.

Geographical classifications are methods to group geographies according to objective criteria, for example classifications based on population density, functional aspects (labour market areas), or geography (mountain areas). Often geographical classifications are based on statistical or administrative geographies to be able to compare statistics between different areas with the same characteristics (for example, urban areas).

Geo-referenced statistical data are data that can be directly presented in space. Geo-referencing, or geospatial referencing, is the process of referencing data against a known geospatial coordinate system, by matching it to known points of reference in the coordinate system, so the data can be processed, queried and analysed with other geographic data.

Geospatial core information is a set of essential geospatial data and services for geocoding other types of information; examples include administrative boundaries, land cover information, addresses, orthophotos/satellite images, transport and hydrographic networks.

Geospatial data are information defined by geometrical boundaries of either administrative or other units that are in geographic information systems (GIS), commonly in the form of polygons.

Grid statistics are spatial statistics geocoded to rectangular grid cells. Each grid cell has the same size and carries a unique code. Ideally the code carries also geocoding information, for example, the lower left corner of the grid cell.

Interoperability is a characteristic of a system (or product) to work with other systems (or products).

Linking defines a process of connecting structured data sources using a system of unique identifiers. While integration describes the process of combining data from different thematic communities, linking refers to technically connecting data in a machine-to-machine environment.

Location is a general term used to describe a place on the surface of the Earth; location data is often used when referring to geospatial information.

The **normalised difference vegetation index (NDVI)** measures how 'green' (in terms of vegetation) an area is. The index is calculated from the reflection of red and near-infrared light. Vegetation appears green because chlorophyll reflects green light; it also absorbs red light. Equally, healthy vegetation reflects near-infrared light. Combining these, healthy vegetation absorbs red light and reflects near-infrared, while unhealthy vegetation does the opposite. The NVDI value is calculated from the percentage reflected values of red and near-infrared light. The calculation is (near-infrared minus red) / (near-infrared plus red). The resulting value ranges from -1 to 1. A value below 0 represents a largely inorganic or dead object. Above 0, the higher the value, the 'greener' the object.

Ontology is a set of concepts and categories in a subject area that shows their properties and the relations between them.

Permanent identifiers (PID) are also known as persistent identifiers. These are long-lasting references to an object.

Point data are those whereby the geographic coordinates are associated with an observation. The value associated with the observation is not of interest, rather it is the location, for example, the point where a disease emerged during an epidemic, or how certain tree species are distributed. Spatial analysis of point data is aimed at quantifying the gap between observations, identifying clusters of data that are more aggregated than if they had been randomly distributed across the territory.

Regional statistics are statistics that are geocoded to administrative and functional geographies.

Spatial analysis or spatial statistics include any of the formal techniques which study entities using their topological, geometric or geographic properties; the phrase refers to a variety of techniques, many still in their early development, using different analytic approaches applied in a wide variety of fields.

Spatial data are statistics which should meet the perception of users in their area of interest (for example their neighbourhood); as such, these data are more detailed than regional statistics. Spatial statistics are geocoded to small administrative or non-administrative geographies.

Spatial unit refers to any set of spatial units that cover a whole territory and are divided into basic (house number, spatial district, statistical district, settlement, municipality, administrative unit) or additional spatial units (local, village or urban community, street, electoral unit), or a grid system (10 km, 1 km, 100 m resolution).

Statistical unit describes one member of a set of entities being studied; this could include persons, households, businesses, buildings or parcels/units of land.

Thematic map is a type of map or chart that is designed to show a particular theme connected with a specific geographic area; these maps can portray physical, social, political, cultural, economic, sociological or agricultural patterns and developments.

Explore further

Methodology

- Inspire Knowledge Base
- Inspire Geoportal
- Handbook of spatial analysis