ECLAC Statistical Briefings

UNITED NATIONS

ECLAC

Global and regional perspectives on the integration of statistical and geospatial information

Contents

The Economic Commission for Latin America and the Caribbean (ECLAC) has joined global efforts to promote the integration of geospatial and statistical information, and understands their value in supporting areas such as decision-making, public policies, national development strategies and the monitoring of the Sustainable Development Goals (SDGs) of the 2030 Agenda for Sustainable Development.

This article aims to provide an overview of the process, including the lines of action and activities that are being pursued at the global level and the measures to put the global proposals into practice in Latin America and the Caribbean.

1. Introduction

Defining public policies that enable the countries of Latin America and the Caribbean achieve full development entails decision-making based on evidence not only from robust data and statistics on sociodemographic, economic and environmental processes, but also from the ability to reference such processes to territories, since they occur in a given place and at a given time. Now more than ever, the information demanded by the global and regional agendas and commitments adopted by the countries of the region require that statistical and geospatial information be integrated into a single data ecosystem that fosters comprehensive information systems and enables analysis of territorial inequalities. ECLAC is working in various areas to promote integration of these systems , with a view to having national and geographical statistical systems that are harmonized and complementary. The following sections outline ongoing initiatives at the international and regional levels that promote such collaborative spaces and encourage the broad participation of Member States of the United Nations.

2. United Nations Expert Group on the Integration of Statistical and Geospatial Information

The United Nations Expert Group on the Integration of Statistical and Geospatial Information (UN EG-ISGI)[1] was established by the Statistical Commission of the United Nations and the Committee of Experts on Global Geospatial Information Management (UN-GGIM) in 2013, to implement a statistical-geospatial framework to support implementation of the 2020 round of population and housing censuses and the 2030 Agenda for Sustainable Development, with the understanding that this framework could be applied to other statistical operations in various areas of development. The aims of the Group of Experts are:

» Provide high-level coordination and a forum for dialogue among representatives of both the statistical and geospatial communities, on global efforts relating to the integration of statistical and geospatial information.

» Play a leadership role by raising awareness and highlighting the importance of reliable, timely, fit-for-purpose, and integrated statistical and geospatial information to support social, economic, environmental and resilience policy decision-making, including at the subnational and regional levels.

» Prioritize and propose work plans and guidelines that advance national and global efforts relating to the integration of statistical and geospatial information, particularly those associated with the Global Statistical Geospatial Framework (GSGF), so that there is increased information to support social, economic, environmental, and resilience policy decision-making, including at the subnational and regional levels.

» Promote and support activities that facilitate the implementation of GSGF, particularly in the international rounds of population censuses and in other censuses, including agriculture censuses and economic censuses, and in global initiatives such as the 2030 Agenda.

» Support the United Nations Statistical Commission and UN-GGIM in the development of norms, principles, guides and standards to significantly increase the availability of high-quality, timely and reliable integrated statistical and geospatial information, including any regional capacity development initiatives.

---

[1] See Expert Group on the Integration of Statistical and Geospatial Information [online] https://ggim.un.org/UNGGIM-expert-group/

3. The Global Statistical Geospatial Framework

The Global Statistical Geospatial Framework (GSGF)[2] enables a range of data to be integrated from both statistical and geospatial communities and, through the application of its five principles (see diagram 1) and supporting key elements, permits the production of harmonized and standardized geospatially enabled statistical data. The resulting data can then be integrated with statistical, geospatial, and other information to inform and facilitate robust data-driven and evidence-based decision-making to support local, subnational, national, regional, and global development priorities and agendas.

» **Diagram 1. Global Statistical Geospatial Framework: principles and key elements**
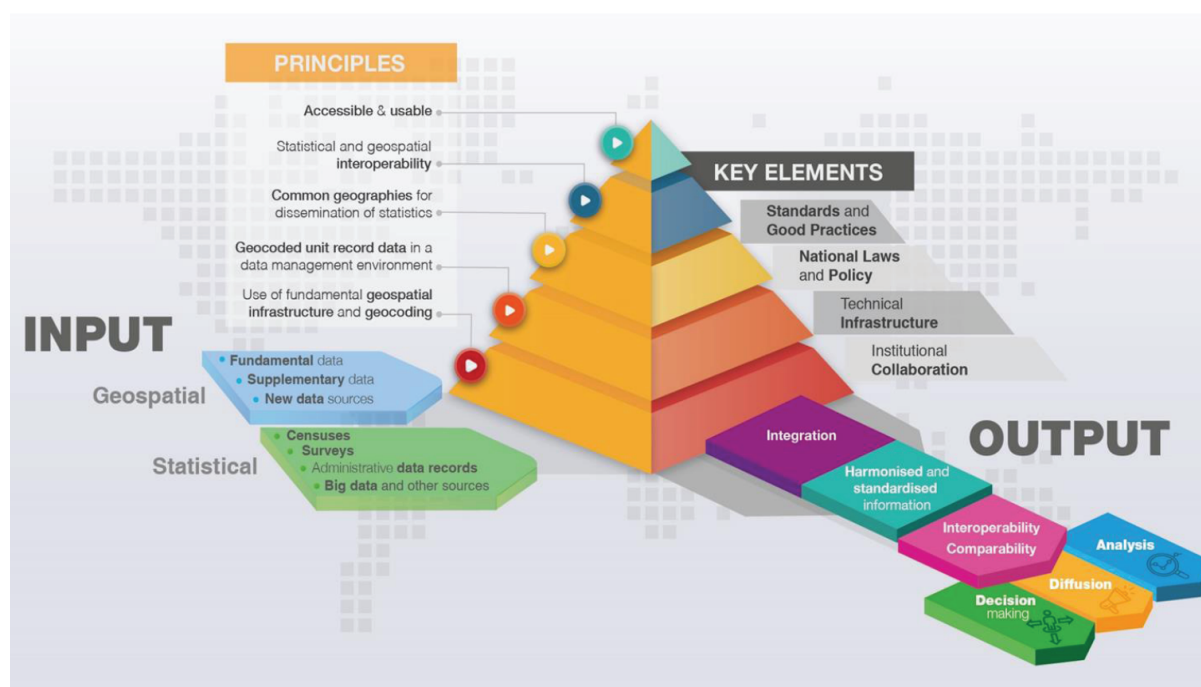
Source: United Nations, *The Global Statistical Geospatial Framework*, New York, 2019 [online] https://unstats.un.org/unsd/statcom/51st-session/documents/The_GSGF-E.pdf.

At its ninth session in August 2019, UN EG-ISGI adopted the GSGF (decision 9/106 on Integration of geospatial, statistical and other related information),[3] as a common method for geospatially enabling statistical and administrative data that ensures data from across a range of sources can be integrated based on location, as well as ensuring that these data can be integrated with other geospatial information. GSGF was subsequently endorsed by the Statistical Commission at its fifty-first session in March 2020 (decision 51/123 on integration of statistical and geospatial information).[4]

---

[2]   See United Nations, *The Global Statistical Geospatial Framework*, New York, 2019 [online] https://unstats.un.org/unsd/statcom/51st-session/documents/The_GSGF-E.pdf.

[3]   United Nations, *Committee of Experts on Global Geospatial Information Management. Report on the ninth session (7–9 August 2019)* (E/2020/46-E/C.20/2019/19), New York, 2019.

[4]   United Nations, *Statistical Commission. Report on the fifty-first session (3-6 March 2020)* (E/2020/24-E/CN.3/2020/37), New York, 2020.

Fundamentally, GSGF enables:

» Integration of data to support the measuring and monitoring of the global indicator framework for the SDGs and the targets of the 2030 Agenda for Sustainable Development and the 2020 round of population and housing censuses.

» Comparisons at local, subnational, national, regional, and global levels for decision-making processes within and between countries and thematic domains.

» Data sharing between institutions, through interoperability of geospatial and statistical information and the development of common tools and applications.

» Unlocking new insights and data relationships that would not have been possible by analysing socioeconomic, environmental or geospatial data in isolation.

» Increased information on smaller geographical areas.

» Increased awareness of methods and tools to assess and manage disclosure risks and to enhance privacy in collection, storage, and dissemination of information.

» Conditions for investment and capability building in geospatial and statistical information.

» Integration of new sources of data to inform the production of high-quality geospatial information, for example Earth observations and other complementary data sources.

» Strengthening of institutional collaboration between the geospatial and statistical communities.

Principle 1: Use of fundamental geospatial infrastructure and geocoding

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. The goal of principle 1 is to obtain high-quality, standardized location references (such as physical addresses, property or building identifiers, or other location descriptions), in order to assign accurate coordinates, small geographic area or standard grid reference to each statistical unit at the microdata/unit record level.

Principle 2: Geocoded unit record data in a data management environment

Principle 2 supports the process of linking or storing high-precision geographic references (which is to say 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. This process applies the address coding infrastructure and fundamental data from principle 1.

Principle 3: Common geographies for dissemination of statistics

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 3 establishes the fundamental importance of balancing existing statistical and administrative geographies with other geographic referencing systems, such as grids, as a basis for establishing common geographies across datasets.

Principle 4: Statistical and geospatial interoperability

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: Accessible and usable geospatially enabled statistics

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 visualize this information seamlessly for geographies of interest. It addresses the need to identify or, where required develop, policies, standards, good practices, and technologies that support these uses.

4. Recommendations for implementing the Global Statistical Geospatial Framework (GSGF) provided by the Integrated Geospatial Information Framework (IGIF)

The IGIF and its Implementation Guide[5] provide guidance for countries that need to determine how to implement GSGF and what steps need to be taken. The steps typically undertaken to integrate geospatial and statistical data are as follows:

1. Develop a strategy: bring focus and provide clear direction to the Geospatial and Statistical Data Integration Initiative through a common vision, mission and goals.

2. Establish a working group: to direct, communicate and oversee the strategy and plan for integrating geospatial information and statistics data, as well as to formalize the relationships between participating agencies through collaborative agreements, and foster an institutional environment for collaboration and cooperation through the implementation of policy, guidelines, standards and common processes.

3. Identify the key stakeholders, such as data producers and users, and decision-makers.

4. Identify the available resources: statistical and administrative data are typically available from national statistics organizations (examples include statistical units, and social, economic, demographic, agricultural, environmental and census statistics). Geospatial information is typically available from national mapping organizations (such as administrative boundaries, addresses, transport, water networks, elevation data, satellite imagery and topographic data).

5. Specify policy, standards, guidelines and norms: policies, standards, guidelines and norms are required to support the utilization, access, analysis and visualization of integrated geospatial and statistical information including terms of use agreements, data release policies, privacy policies, data standards to enable interoperability and guidelines for the security of information.

6. Develop the methodology: Develop the methodologies and procedures that standardize the collection, generation and maintenance of data, both geographic and tabular:

    »   Address: use a standard for the collection and assignment of addresses nationwide to effectively capture the address label and physical location.

    »   Coding: establish coding classifications for disaggregation into spatial entities, such as government administrative boundaries or census areas.

---

[5]   Expert Group on the Integration of Statistical and Geospatial Information, "Appendices. Strategic Pathway 4 Data", Integrated Geospatial Information Framework [online] https://ggim.un.org/IGIF/documents/SP4-Appendices-26Feb2020-GLOBAL-CONSULTATION.pdf.

» Verification: implement address verification at point of address data entry (for computer-based capture/ mobile applications) to better manage and maintain data quality and reduce time spent on data cleansing.

» Georeferencing: align geographic data to a known coordinate system so it can be analysed, viewed and queried with other data.Geocoding: adopt common geocoding practices at the national level.

» Quality: specify a quality model and measures for both geospatial and statistical data.

» Retrieval: establish an effective method for storing and managing integrated geospatial and statistical data files so that they can be automatically updated and retrieved without the need for reprocessing.

» Linking: evaluate open-source, bespoke and commercial off-the-shelf (COTS) services for connecting data, with consideration to the Open Geospatial Consortium (OGC) Table Joining Services standard.

» Make data accessible: use standard services (such as Web Map Service (WMS) and Web Feature Service (WFS)) to serve data and distribute results over the Web.

5. The UN EG-ISGI 2020-2022 work plan for integration of statistical and geospatial information

UN EG-ISGI of UN-GGIM has developed a work plan[6] containing five activities supported by the formation of task teams, as described below:

a. Support the development of IGIF:

Develop a matrix that maps the interlinkages between the principles and key elements of GSGF and the Strategic Pathways of IGIF to further document the relationship between GSGF and IGIF and develop documentation that enables the mutual implementation of both frameworks.

b. Develop a communications strategy:

Develop a communications strategy that enables: promotion and awareness-raising of the outputs of UN EG-ISGI, to assist leaders and change agents; guidance to support the 2020 round of population censuses; development of communication material that can be used to identify the differences between GSGF and IGIF; and guidance to assist with the reporting of progress towards the attainment of the SDGs.

c. Establishment of task teams

» *Task Team on Capacity-Building*

The aim of this team is to develop guidance regarding the implementation and strengthening of statistical-geospatial capacity in countries, through a stocktaking of activities in statistical-geospatial integration that allows for mitigation of potential duplication (or divergence) of work.

» *Task Team on Privacy and Confidentiality*

The aim of this team is to provide a document that develops guidance and recommendations for the management of privacy and confidentiality issues when using, sharing or releasing geospatial data within national statistical offices (NSOs).

---

6   See "Expert Group on the Integration of Statistical and Geospatial Information Work Plan 2020 - 2022" [online] https://ggim.un.org/ meetings/GGIM-committee/10th-Session/documents/EG-ISGI_Work%20Plan_2020-2022.pdf.

» *Task Team on the Principles of the GSGF*

The aim of this team is to help provide guidance to enable countries to implement GSGF and to help develop the substance of GSGF, as a living document. This will be achieved through three work streams: geocoding, common geographies, and interoperability.

d. Consider mechanisms to raise engagement and participation within the work programme of UN EG-ISGI by Member States

Coordinate with UN-GGIM Europe (and identify complementary work ongoing within the regions) on their work on the integration of statistical and geospatial information. Identify potential interlinkages with other groups (including the Inter-Agency and Expert Group on Sustainable Development Goal Indicators Working Group on Geospatial Information (IAEG-SDGs: WGGI), the Open Geospatial Consortium's Statistical Domain Working Group, and the UN-GGIM Working Group on Legal and Policy Frameworks for Geospatial Information Management) to support collaboration on this work programme and its delivery.

e. Illuminate case studies and examples of good practice in the adoption of GSGF

Seek and collate use cases that can be used to support the implementation of GSGF at the national and regional level. Seek and collate use cases that demonstrate how GSGF has been used to benefit the national or regional response to the coronavirus disease (COVID-19) pandemic, or could have been used if there were available.

6. Activities being carried out in the Americas to support the integration of statistical and geospatial information

a. The MEGA project

The Statistical and Geospatial Framework for the Americas (MEGA)[7] project is being developed under the auspices of the Regional Committee of United Nations Global Geospatial Information Management for the Americas (UN-GGIM: Americas) for the integration of statistical and geospatial information.

MEGA defines an information infrastructure made up of statistical and geospatial information which must be connected and conceptually integrated, which is to say it connects geospatial information with statistics that describe socioeconomic and environmental attributes. MEGA will allow the integration of statistical and geospatial information in a unified way for the region with common principles and standards, strengthening dissemination for decision-making. MEGA is fully aligned with the five principles of GSGF, developed by UN EG-IGSI. Version 1.0 of MEGA consists of population data by sex and housing data, and establishes three levels of territorial disaggregation.

b. The Central America project

The project on Integration of Statistical and Geospatial Information in Central America[8] is financed by the Pan-American Institute of Geography and History (PAIGH), led by the United States Census Bureau and by the National Institute of Statistics and Geography of Mexico (INEGI) within the framework of UN-GGIM:Americas, with the collaboration of ECLAC and the National Centre of Geographical Information of Spain (CNIG). It works with the seven Central American countries: Belize, Costa Rica, El Salvador, Guatemala, Honduras, Nicaragua and Panama.

---

[7] See Regional Committee on Global Geospatial Information Management for the Americas, "Mega Project" [online] http://www.un-ggim-americas.org/en/assets/modulos/proyectos.html?proyecto=1.

[8] See Regional Committee on Global Geospatial Information Management for the Americas, "Central America Project" [online] http://www.un-ggim-americas.org/en/assets/modulos/proyectos.html?proyecto=4.

The project applied in 2018 to the PAIGH Technical Assistance Program for 2020 and was approved on 23 October 2018 during the forty-eighth meeting of the PAIGH Directing Council that took place in Santa Cruz, Bolivia. It will end in 2021.

The objectives of the project include the creation of a geospatial platform to enable collaborative statistical and geospatial integration work in the region and a technical workshop to identify innovative ways of visualizing and analysing integrated statistical and geospatial information.

Current project work includes collaboration on georeferenced health data to support decision-making relating to the COVID-19 pandemic.

c. **Project on innovative approaches to examining inequality, by integrating different data sources in Latin America and the Caribbean.**

This project, implemented by ECLAC, is expected to increase the capacities of selected countries to integrate geospatial and statistical data and information by adopting regional and global frameworks, standards and methodologies for evidence-based policies, enabling them to address the 2030 Agenda and generate disaggregated data to support management of COVID-19 with a territorial perspective.

The key activities to be carried out as part of this project are:

» Assessment of capacity development needs and formulation of the training programmes needed to strengthen the technical capabilities of national geospatial information agencies and NSOs.

» Technical assistance missions to national geospatial information agencies and NSOs to facilitate knowledge-sharing, exchange best practices, assess options and formulate and implement nationally developed approaches in GSGF, with emphasis on the use of geostatistical data, supporting COVID-19 management and vulnerability analysis in the integration of geospatial and statistical data and information.

» National and regional workshops with representatives of national geospatial information agencies, NSOs and other stakeholders to promote coordination and establish roadmaps that integrate, disseminate and share geospatially enabled statistical data.

d. **Project on geospatial capacity-building for national statistical offices for follow-up of the 2030 Agenda in Latin America and the Caribbean**

This project, implemented by ECLAC with the support of the European Union, aims to strengthen capacities for the design, construction and implementation of statistical geoportals in NSOs, to manage, integrate and visualize geospatially enabled statistical data.

Such tools will not only improve internal processes at NSOs, but will also provide capabilities to integrate statistical data with other geospatial information and Earth observation data and sources, enriching national data ecosystems to support decision-making, sustainable public policy and subnational and local development.

The project will also contribute to strengthening statistical databases by improving their interchangeability capabilities under international standards and regional benchmarks, consolidating regional assets such as CEPALSTAT that contribute to comparable regional analysis on sustainable development in the region.

This project was begun in a transitional phase in which NSOs are making significant efforts to close the digital divide, seeking to collectively modernize their statistical outputs, systems and processes, transform their operations and develop new relevant metrics and indicators for statistical purposes. Crucially, this includes introducing standards-based, metadata-driven processes and infrastructure.

The project has five components: diagnostic assessment of the current situation in NSOs at the regional level and in selected countries of the European Union, relating to the use of geospatial information in their production and dissemination processes; design of a prototype statistical geoportal to be implemented in each of the target countries of the project, to store, process and publish online (as map services) geocoded statistical data; development of the component applications of the software architecture and functional design of the geoportal, incorporating quality assurance activities; implementation of the statistical geoportal prototype, including software installation and integration with the map service, metadata and other public data at the national level; and operation of the statistical geoportal in the target countries.

ECLAC

Economic Commission for Latin America and the Caribbean (ECLAC)
Comisión Económica para América Latina y el Caribe (CEPAL)
www.eclac.org