Global Statistical Geospatial Framework: Linking Statistics and Place

Current status and plans for development
July 2018

United Nations Expert Group on the

Integration of Statistical and Geospatial Information

Contents

Executive Summary

This paper presents the current status and plans for development of the Global Statistical Geospatial Framework. The five high-level principles in the Framework have been endorsed by the international statistical and geospatial communities – as represented by the UN Statistical Commission (UNSC) and United Nations Committee of Experts on Global Geospatial Information Management (UN-GGIM).

The UN Expert Group on the Integration of Statistical and Geospatial Information (UN EG-ISGI) was established by the UNSC and UN-GGIM with a mandate to develop an international statistical geospatial framework. The Global Forum on the Integration of Statistical and Geospatial Information, convened in New York in August 2014, identified that "there is an urgent need for a mechanism, such as a global statistical-spatial framework, to facilitate consistent production and integration approaches for geo-statistical information."[1] The resulting Global Statistical Geospatial Framework is a principles based Framework that has evolved from Australia's Statistical Spatial Framework and has been guided by a global consultation process. The UN EG-ISGI is continuing to pursue areas of further detail to permit consolidation and implementation of the Global Framework. Once complete, this detail will be brought to the UNSC and UN-GGIM for consideration.

The Global Statistical Geospatial Framework provides 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. This will enable:

- new, better and more integrated information for analysis and decision making processes;
- comparisons within and between countries in a more harmonised manner;
- increased information on smaller geographic areas[2];
- the development of common tools/applications to support the integration and sharing of data;
- commercial development of geospatial tools that will further support data integration; and,
- generally more efficient production of information.

The five high-level principles of the Global Statistical Geospatial Framework

Principle 1: Use of fundamental geospatial infrastructure and geocoding.

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

Principle 3: Common geographies for dissemination of statistics.

Principle 4: Statistical and geospatial interoperability – Data, Standards and Processes.

Principle 5: Accessible and usable geospatially enabled statistics.

[1] http://ggim.un.org/meetings/2014-Global_Forum/documents/Summary-Report%20of%20the%20Global%20Forum.pdf

[2] *For the purpose of this paper, the terms "geographies" and "geographic areas" will be used to describe a broad range of geographic areas or regions that define places, from small to large areas.*

The Global Statistical Geospatial Framework comes at an important time as statistical and geospatial agencies work to modernise and transform their models of operation and infrastructure. The Global Framework will also be critical to support the work occurring on the 2020 Round of Population Censuses and the 2030 Agenda for Sustainable Development.

The usefulness of these high-level principles and Global Framework has already been demonstrated by the adoption of these principles by a number of countries, each at different starting points in their statistical and geospatial infrastructure development (e.g. Sweden, Mexico, New Zealand, Egypt and South Africa), as well as strong interest from regions, such as Europe, Latin America and the Arab States. Many member states responding to the international consultation also identified their desire to implement the Framework within their national circumstances.

These high-level principles reflect best practice in the statistical and geospatial arenas and their development has strengthened the partnerships between these communities. The endorsement of the Global Statistical Geospatial Framework principles represents a significant achievement by the UN-GGIM in the early years of its functioning.

Global Statistical Geospatial Framework – Full Paper

Background

The Statistics Division of the Department of Economic and Social Affairs of the Secretariat established the Expert Group on the Integration of Statistical and Geospatial Information in 2013, comprising members of both the statistical and geospatial professional communities from Member States. The Expert Group reports to both the United Nations Statistical Commission (UNSC) and the United Nations Committee of Experts on Global Geospatial Information Management (UN-GGIM) at each of their respective annual sessions.

The overall objectives and functions of the Expert Group are to pursue the implementation of the Global Statistical Geospatial Framework and its application towards the 2020 Round of Censuses with the understanding it would apply to other initiatives including other censuses, such as agriculture and economic censuses, and the implementation of the 2030 Agenda for Sustainable Development.

A proposal was developed based on Australia's Statistical Spatial Framework (SSF) which was sent out for global consultation and as a result of the consultation further refinements were made.

In August 2016, the 6th Committee of Experts on Global Geospatial Information Management adopted the five guiding principles of the Framework. Then in March 2017, the 48th session of the Statistical Commission endorsed the 5 guiding principles of the Global Statistical Geospatial Framework. This UNSC meeting also extended the mandate of the Expert Group to "become the overall coordination group for all activities in the area of the integration of statistical and geospatial information".

Context

It is now readily accepted that integrating statistical and geospatial information is critical for:

- local, sub-national, national, regional, and global decision making processes;
- measuring and monitoring the targets and global indicator framework for the Sustainable Development Goals (SDGs) of the 2030 Agenda for Sustainable Development;
- supporting data sharing between institutions and enhancing the interoperability of geospatial and statistical information;
- unlocking new insights and data relationships that would not have been possible by analysing socio-economic, environmental or geospatial data in isolation;
- promoting investment and capability building in geospatial and statistical information;
- building institutional collaboration between geospatial and statistical communities; and,
- examining new sources of data that includes geospatial information, for example mobile phone data.

The challenge now is “how best to achieve this integration in an effective and consistent way”[3]. The Global Forum on the Integration of Statistical and Geospatial Information, convened in New York in August 2014, identified that “there is an urgent need for a mechanism, such as a global statistical-spatial framework, to facilitate consistent production and integration approaches for geo-statistical information.”[4]

The Expert Group considers the best way to achieve consistent integration is through having a common method of geospatially enabling statistical and administrative data and integrating this with geospatial information through an internationally agreed framework.

This will enable:

- new, better and more integrated information for analysis and decision making processes;
- comparisons within and between countries in a more harmonised manner;
- increased information on smaller geographic areas[5];
- the development of common tools/applications to support the integration and sharing of data;
- commercial development of geospatial tools that will further support data integration; and,
- generally more efficient production of information.

For National Statistical Organisations (NSOs) efforts to improve the integration of statistical and geospatial information occurs in an environment where NSOs are seeking to collectively modernise their statistical productions, systems and processes, to transform their operations and to derive new relevant metrics and indicators for statistical purposes. Critically this includes the introduction of standards based, metadata driven infrastructure and processes.

For National Geospatial and Mapping Agencies (NGMAS) this work occurs at a time where there is a collective effort and will to enhance the management and use of geospatial information. This includes efforts to specify the fundamental or core geospatial datasets that are required to support geospatial activities within nations and for international efforts; for example, to support reporting against the Sustainable Development Goals. These fundamental datasets are part of national geospatial infrastructures, which supports the geocoding of statistics and includes administrative and statistical geographies that support the integration of statistical information with geospatial information.

[3] http://unstats.un.org/unsd/statcom/doc13/2013-2-ProgReview-E.pdf

[4] http://ggim.un.org/meetings/2014-Global_Forum/documents/Summary-Report%20of%20the%20Global%20Forum.pdf

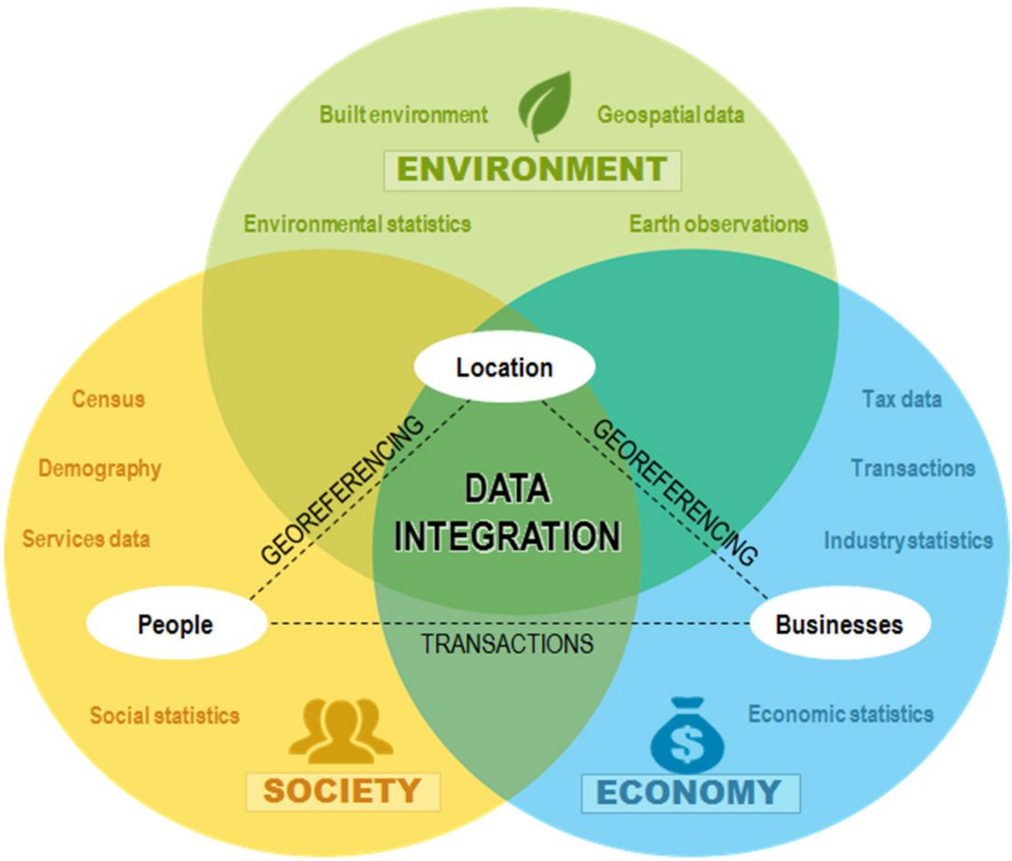
[5] *For the purpose of this paper, the words and terms “geographies” and “geographic areas” will be used to generically describe a broad range of geographic areas or regions that define places, from small to large areas.*

The Global Statistical Geospatial Framework

The international statistical and geospatial communities have recognised the challenge of better integration of geospatial and statistical information and have responded by mandating the United Nations Expert Group on the Integration of Statistical and Geospatial Information (UN EG-ISGI) to develop a Global Statistical Geospatial Framework. The Global Framework acts as a bridge between statistics and geospatial information, between NSOs and NGMAs, and between statistical and geospatial standards, methods, workflows, and tools.

The Global Framework provides the international community with a common approach to connecting socio-economic and environmental data through location, and improves the accessibility and usability of this geospatially-enabled data. Figure 1 highlights the importance of location information as an integrating tool between the three broad data domains: society, the economy and the environment.

Figure 1: Location as a link between society, the economy and the environment



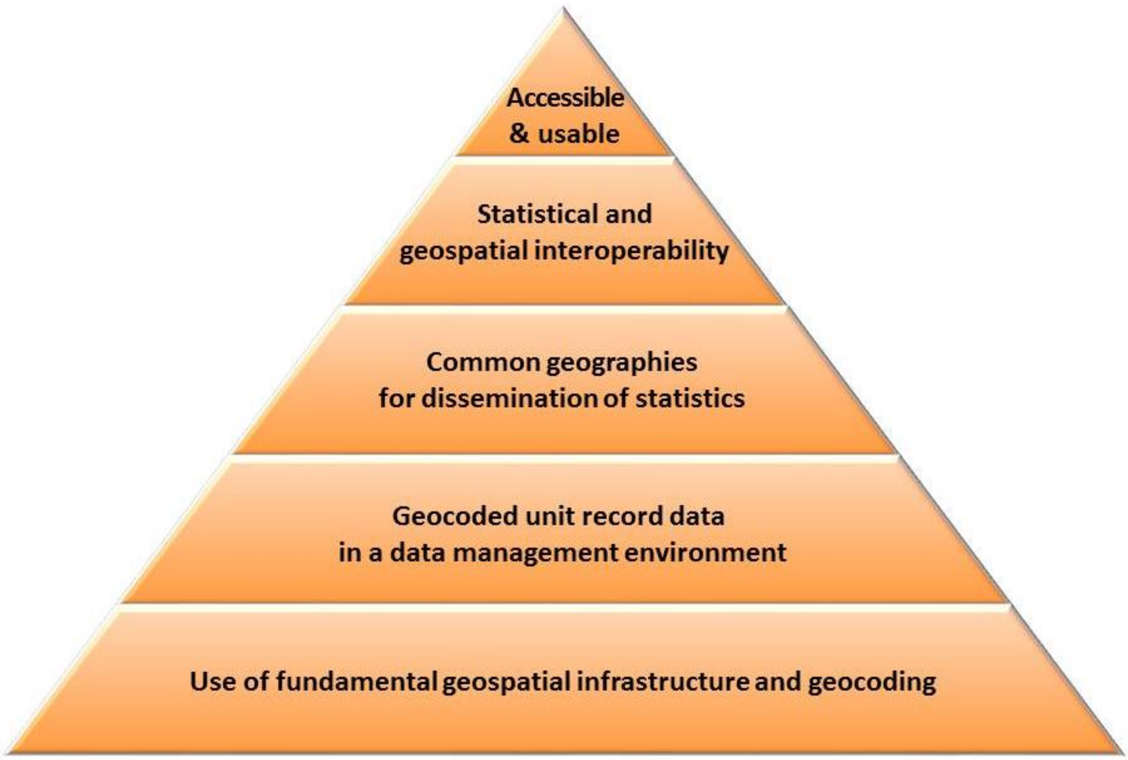
At this point in time, the Global Statistical Geospatial Framework focuses on the socio-economic and environmental statistical data traditionally produced by NSOs. The UN EG-ISGI will monitor the scope of the Framework with a review point in 2019. The intention of the Expert Group is for the Framework to be inclusive of all statistical and geospatial data, and to enable and encourage NSOs to look beyond traditional data sources and methods.

The Global Framework also provides an important communication tool and a common platform for the international community to discuss and understand the geospatial capabilities requirements for statistical information. This has been a vital element in the journey towards integrating geospatial capability in statistical transformation and modernisation efforts, and in development efforts in many national statistical systems. It has also provided the mechanism to connect statistical information into the efforts to improve geospatial information management globally, regionally and nationally.

What is the Global Statistical Geospatial Framework?

The Global Statistical Geospatial Framework is a high-level framework that consists of five broad principles that are considered essential for integrating geospatial and statistical information (see the orange layers in Figure 2 below).

Figure 2: Global Statistical Geospatial Framework



The Global Statistical Geospatial Framework Principles – Goals and Objectives

Each of the high-level principles in the Global Statistical Geospatial Framework are defined by a set of goals and objectives, and are supported by international, regional and applicable domestic standards and best practice. These principles and the associated goals and objectives are discussed in detail in the following sections of this paper. The current supporting standards and best practice that will support countries implementing the Global Framework are listed for information in Appendix 2.

These goals and objectives, as well as the supporting standards and best practice that will provide further detailed guidance for countries implementing the Global Framework, are being further developed by the Expert Group. These will be brought to the UNSC and UN-GGIM for consideration when finalised and published.

Where standards, policies or datasets required to support the Framework do not currently exist, the Global Framework provides a clear mandate for their establishment. Collaboration between countries and within the United Nations Expert Group on the Integration of Statistical and Geospatial Information (UN EG ISGI) provides a mechanism to assist with the formation and establishment of these standards, policies or fundamental datasets – both within member states and internationally.

Case studies from Australia and a range of other countries and regions, included in Appendix 3, show how international, regional and domestic standards can be applied in the context of this Framework.

Principle 1: Use of fundamental geospatial infrastructure and geocoding

The Global Framework requires a common and consistent approach to establishing the location and a geocode for each unit in a dataset, such as a person, household, business, building or parcel/unit of land. A corresponding record of the relevant time or date for each instance of location information recorded should also be associated with each unit record.

The goal of this principle is to obtain a high quality, standardised physical address, property or building identifier, or other location description, in order to assign accurate coordinates and/or a small geographic area or standard grid reference to each statistical unit (i.e. at the microdata level). Time and date stamping these locations will place the unit both in time and in space. An alternate approach to geocoding for recording location is to use direct or indirect capture of coordinates (e.g. from Global Navigation Satellite Systems (GNSS) and maps respectively) from field work. Where this level of precision is not possible using current geospatial and statistical infrastructure within a country, adaptations using more general location descriptions and/or larger geographies will be necessary.

The Expert Group notes that geocoding statistical units using point referencing is highly preferable when compared to only associating statistical units with a geographic region (i.e. a polygon). The use of point referencing allows for considerable adaptability to changes in geographic regions over time or to adapt to new geographies that emerge. Where an established point based geospatial infrastructure (such as a georeferenced address or building register) does not exist in a country, it is recommended countries test the implementation of other point-based referencing for unit record data (for example using GPS).

The process of obtaining locations and geocodes should use relevant, fundamental geospatial data[6] from National Spatial Data Infrastructures (NSDI) or other nationally agreed sources. These processes are generally referred to as geocoding, which is defined in detail in Appendix 1.

[6] http://ggim.un.org/meetings/GGIM-committee/documents/GGIM5/E-C20-2015-4%20Fundamental%20Data%20Themes%20Report.pdf

To ensure that all statistical data are consistently geospatially enabled using these methods the following objectives should be met:

- Address, property, building, and location information are accurate and consistent, meeting nationally agreed standards.
- Geocoding results are as accurate and consistent as possible.
- Consistent management of any geocoding issues through application of standardised approaches.

The Expert Group have identified a range of international and national standards, frameworks, infrastructure, and best practice that are relevant to this principle. Their potential uses in support of this principle are outlined in Appendix 2.

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

The Global Framework recommends that the linkage of a geocode for each statistical unit record in a dataset (i.e. a person, household, business, building or parcel/unit of land) occurs within a data management environment that is based on agreed good practice principles. Persistent storage of a high precision geocode enables any geographic context to be applied when preparing the data for release in the future (i.e. in aggregating data into a variety of larger geographic units or to adapt to changes in geographies over time). Moreover, geocodes can enable data linking processes that aim to integrate information of varying nature and sources, such as environmental information.

This component of the Global Framework also recommends that established data management tools, techniques and standards be used to facilitate the integration and management of the geocode with the statistical dataset, including address to geocode linking mechanisms. This will ensure that all statistical data is consistently geospatially enabled and that the following objectives are met:

- Consistent and interpretable geocode information.
- Simplified aggregation of data for larger geographies through storage of an identifier or code for a small area geography or standard grid square for each unit record.
- Adaptation to changes to existing geographies or to allow compilation of data for new geographies (while also being conscious of the consequential confidentiality risks associated with differencing).
- Enable the flexible use of geocoded unit records in future analysis and visualisation.
- Effective data management, including protection of privacy and compilation of metadata.
- Clear maintenance and custodianship roles.

The Expert Group have identified a range of international and national standards, frameworks, infrastructure, and best practice that are relevant to this principle. Their potential uses in support of this principle are outlined in Appendix 2.

Principle 3: Common geographies for dissemination of statistics

To enable comparisons across datasets from different sources, the Global Statistical Geospatial Framework recommends that a common set of geographies be used for the display, reporting and analysis of social, economic and environmental information.

The UN EG-ISGI recognises the fundamental importance of traditional statistical and administrative geographies as a basis for the common geographies recommended in this principle.

The Expert Group also recommends NSOs consider the benefits of gridded data. Gridded data can be both a rich source of information and a consistent geography for disseminating and integrating information. Recent global efforts have culminated in the development of a Discrete Global Grid Systems (DGGS) standard which has been developed under the auspices of Open Geospatial Consortium (OGC). This System offers further options in the use of grids within the context of the principle of common geographies and in geospatially enabled statistics.

Use of a common set of geographies will ensure that all statistical data is consistently geospatially enabled and that users can discover, access, integrate, analyse and visualise statistical information seamlessly for geographies of interest. This will allow the following objectives to be met:

- Data from disparate sources can be integrated using common geography.
- Visualisation and analysis is simplified.
- Metadata supports data integration and use.
- Conversion of data between geographies is supported, through standard conversion mechanisms (e.g. correspondences[7]).

The Expert Group have identified a range of international and national standards, frameworks, infrastructure, and best practice that are relevant to this principle. Their potential uses in support of this principle are outlined in Appendix 2.

Principle 4: Statistical and geospatial interoperability – Data, Standards and Processes

Both the statistical and geospatial data communities operate their own general data models and metadata capabilities; however, these are often not universally applied. The statistical community use the Generic Statistical Information Model (GSIM), the Statistical Data and Metadata Exchange (SDMX), and Data Documentation Initiative (DDI) mechanisms. The geospatial community use the General Feature Model (GFM) and developed the ISO19115 metadata standard, plus a number of application specific standards.[8]

---

[7] For more information on correspondence methods see: http://www.abs.gov.au/websitedbs/D3310114.nsf/home/Correspondences

[8] For a discussion on these statistical and geospatial models and metadata standards, see: http://ggim.un.org/meetings/2015-2nd_Mtg_EG-ISGI-Portugal/documents/Connecting%20Geographic%20and%20Statistical%20Information%20Standards%20EG-

The Expert Group recognises the benefits of greater interoperability between statistical and geospatial data and metadata standards, from cataloguing to data interchange. Overcoming structural and syntactic barriers between data and metadata from different communities and providers will enhance the efficiency of discovery, access, and use of geospatially enabled data. The Expert Group has noted that a number of initiatives have started internationally on these topics and encourages these efforts, while emphasising the importance of harmonisation across this work.

Within the statistical community there is a need to build geospatial processes and standards into statistical business processes in a more consistent manner. The Expert Group has recognised that, to ensure this occurs, a top down approach of incorporating geospatial frameworks, standards and processes more explicitly into the Common Statistical Production Architecture (CSPA) and its components is required. In particular, the General Statistical Business Process Model (GSBPM) needs to make greater reference to the use of geospatial data and methods in the statistical production process, particularly the data, standards and methods that are incorporated into the Global Statistical Geospatial Framework. Eurostat through the GEOSTAT grant projects has funded the European Forum for Geography and Statistics to undertake research in this area.[9] The High-Level Group for the Modernisation of Official Statistics is currently reviewing GSBPM and is considering how to incorporate geospatial processes and standards into statistical business processes.

By encouraging greater interoperability of statistical and geospatial data, standards and processes within the context of the Global Statistical Geospatial Framework the following objectives will be met:

- Greater efficiency and simplification of the creation, discovery, integration and use of geospatially enabled statistics and geospatial data.
- Implementation of service based or machine-readable access mechanisms (e.g. APIs) that provide greater efficiency of access and use, and allow adaptation and evolution of uses through time.
- Increasing the potential application of a larger range of data and technologies.
- A wider range of data available and accessible for use in comparisons and analysis in decision making.

The Expert Group have identified a range of international and national standards that are relevant to this principle. Their potential uses in support of this principle are outlined in Appendix 2.

ISGI%202015.pdf and http://ggim.un.org/meetings/2015-2nd_Mtg_EG-ISGI-Portugal/documents/Metadata%20interoperability%20cover%20paper%20EG-ISGI%202015.pdf

[9] https://www.efgs.info/geostat/geostat-3/

Principle 5: Accessible and usable geospatially enabled statistics

This component of the Global Framework emphasises the need to identify or, where required, develop policies, standards and guidelines that support the release, access, analysis and visualisation of geospatially enabled information.

These policies, standards and guidelines will highlight the wide range of legislative and operational issues that organisations need to be aware of when releasing and analysing information about people and businesses. One important aspect of this principle is to ensure data can be accessed using safe mechanisms that protect privacy and confidentiality but also enable access to data in order to undertake analysis that informs decision making. Other issues of relevance include: data quality in its different dimensions (particularly with regard to reliability, timeliness, and relevance), analysis, dissemination and visualisation.

This principle will ensure that custodians release data in appropriate forms and users can discover, access, integrate, analyse and visualise statistical information seamlessly for geographies of interest, thereby meeting the following objectives:

- Data custodians can release data with confidence, with privacy and confidentiality protected.
- Data users can discover and access geospatially enabled statistics.
- Data users can undertake analysis and visualisation.
- Web services enable machine-to-machine access, as well as dynamic linkage of information.

The Expert Group have identified a range of international and national standards, frameworks, infrastructure, and best practice that are relevant to this principle. Their potential uses in support of this principle are outlined in Appendix 2.

Future plans to develop the GSGF

Topics for further work

The Expert Group recognises that there are a number of areas of detail that come under each of the five high-level principles of the Global Framework, where further work is required. Pursuing global cooperation on these topics will ensure that:

- maximum benefit from the Global Statistical Geospatial Framework is realised;
- implementation of the Global Framework within countries and regions is simplified;
- fundamental national and international challenges in this area are addressed;
- knowledge and capability is shared globally; and
- collaboration and communication between the statistical and geospatial communities continues to be promoted.

The Expert Group have agreed on the following areas of detail that require further work:

- Application of the principles of the Framework to the development of statistical and geospatial information to support the global indicator framework for Sustainable Development Goals (SDGs).
- Build capability through the application of the Global Statistical Geospatial Framework and geospatial technologies to the 2020 Round of Population Censuses;
- Work to operationalise the principles of the Framework to ensure they are implemented and consolidated;
- Enhance collaboration and partnership between statistical and geospatial organisations;
- Work towards consistent terminology internationally and across communities;
- Protect confidentiality within statistics released for small geographic areas and across different geographies;
- Ensure data is interoperable between statistical and geospatial domains through connecting, extending and enhancing information (data and metadata) standards and information architectures (i.e. the Common Statistical Production Architecture (CSPA) and the General Statistical Business Process Model(GSBPM)), and the development and application of linked data methods;
- Investigate the application of statistical, administrative and grid geographies for data release and any issues associated with managing confidentiality and data comparability;
- Develop and share methods for ensuring effective and authoritative geocoding; and
- Contribute to the broader discussion on the use of Big Data in official statistics and geospatial information; for example, The United Nations Statistics Division Global Working Group on Big Data for Official Statistics[10].

[10] http://unstats.un.org/unsd/bigdata/

Appendix 1: Definitions

Geocoding

- For the purposes of the Global Statistical Geospatial Framework, geocoding is generally defined as the process of geospatially enabling statistical unit records so that they can be used in geospatial analysis.
- More specifically, geocoding is the process of *linking* unreferenced *location information* (e.g. an address), that is associated with a *statistical unit*, to a *geocode* (i.e. a geospatially referenced object); alternatively, the geocode can be directly incorporated into the statistical unit record.

    - *Geocodes* are, preferably, fine scale geospatially referenced objects that are stored as a geometry data type, such as: location coordinates (i.e. x, y, z coordinates) and/or small area geographies (e.g. mesh blocks, block faces or similar small building block geographies). Larger geographic units, such as enumeration geographies, can be used as geocodes where finer scale geospatial units are not available.

    - *Location information* can include addresses, property or building identifiers, as well as other location descriptions, such as enumeration geographies and other standardised (e.g. what3words reference) and non-standardised (e.g. village names) textual descriptions of a location.

    - *Statistical units* can include persons, households and living quarters, businesses, buildings or parcels/units of land.

    - *Linkage* of a geocode to a statistical unit record can occur through use of standard geographic coding systems, a Uniform Resource Identifier (URI) or through other computer based linkage mechanisms.

Georeferencing

- Georeferencing is a set of broad processes that includes *geocoding*. The following definition is included to aid understanding between the two terms.

Georeferencing, or geospatial referencing, is the process of referencing data against a known geospatial coordinate system, by matching to known points of reference in the coordinate system (e.g. image rectification to survey points or addresses linked to parcel centroids), so that the data can be viewed, processed, queried and analysed with other geographic data.

Appendix 2: Relevant international and national standards, frameworks, infrastructure and best practice for the five Global Framework Principles.

Principle 1: Use of fundamental geospatial infrastructure and geocoding

- The Global Geodetic Reference Frame (GGRF), and aligned regional or national Geodetic Reference Frames.

- Fundamental geospatial data from the relevant National Spatial Data Infrastructure (NSDI) or other equivalent, nationally agreed geospatial data sources; in accordance with the Statement of Shared Guiding Principles for Geospatial Information Management[11].

- Nationally consistent address collection standards, to effectively capture the physical address; or location reporting standards, where standard addressing systems do not exist.

- Infrastructure that supports point-of-entry address or location validation (for computer and internet based capture), which will improve the quality of address and location information capture and the quality of the geocoding outputs. It will also reduce the time spent correcting addresses and locations provided by respondents in Principle 2.

- Adopting common and practical methods of address and location data capture as a key element of improving the quality of address and location information in administrative and collection based datasets.

- Implement common geocoding practices for statistical and administrative information; for example, use of common methods and infrastructure, such as national and regional geospatial reference systems based on fundamental or National Spatial Data Infrastructure datasets, including the use of national registers. It is expected that in many contexts these common practices may need to be developed.

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

- Statistical and geospatial data management frameworks based on agreed good practice principles, and preferably internationally accepted

- Geocoding metadata standards and application guides, including point-of-entry address or location validation and geocoding

- National privacy laws and/or agreed privacy standards, in accordance with the United National Fundamental Principles for Official Statistics (UN FPOS)

- Common geographic classifications

- Geospatial data management, including use of Global or national/regional Geodetic Reference Frames

[11] http://ggim.un.org/documents/statement%20of%20shared%20guiding%20principles%20flyer.pdf

Principle 3: Common geographies for dissemination of statistics

- Common geographic classification, including use of administrative and statistical geographies that are complemented by use of grid type geographies.
- Standards or guidance on the use of geographies for dissemination of data.
- International statistical and geospatial metadata standards.
- Systems and methods to correspond data between geographies.[12]
- National privacy laws and/or agreed privacy principles, in accordance with the United National Fundamental Principles for Official Statistics (UN FPOS).[13]

Principle 4: Statistical and Geospatial Interoperability – Data, Standards and Processes

- As part of the Open Data initiative, the World Wide Web Consortium (W3C) has suggested Data Catalog Vocabulary (DCAT) as a standard supporting the discovery use case of all types of information. Specific application profiles of DCAT to geospatial information (GeoDCAT) and statistics (StatDCAT) are being developed which are interoperable with ISO19115 and SDMX respectively.[14]
- The OGC Table Joining Service (TJC) and INSPIRE directive also promote interoperable and metadata standards.

Principle 5: Accessible and usable geospatially enabled statistics

- Where feasible adopt policies that maximise access to and use of open, free and unrestrictive geospatial information for innovation, efficient and effective decision making and a geospatially enabled society.[15]
- Standards or guidance on the use of geographies for dissemination, visualisation and analysis.
- Access to machine-readable data through web services (APIs, open service standards, etc.).
- National privacy laws and/or agreed privacy standards, in accordance with the United Nations Fundamental Principles for Official Statistics (UN FPOS).

---

[12] For more information on correspondence methods see: http://www.abs.gov.au/websitedbs/D3310114.nsf/home/Correspondences

[13] http://unstats.un.org/unsd/dnss/gp/fundprinciples.aspx

[14] http://ggim.un.org/meetings/2015-2nd_Mtg_EG-ISGI-Portugal/documents/UN-GGIM%20EG%20Lisbon%20meeting%20Session%206%20background%20paper%20Metadata%20DCAT.pdf

[15] http://ggim.un.org/ggim_20171012/docs/meetings/GGIM5/E-C20-2015-10%20Statement%20of%20Shared%20Principles%20Report.pdf

Appendix 3: Applying the Global Statistical Geospatial Framework

The principles of the Global Statistical Geospatial Framework are high level and flexible enough that they can be adapted and applied to a wide variety of country or regional contexts. By keeping the principles broad they can then be adapted to local circumstances, while still encouraging the use of international standards and methods.

Australia's Statistical Spatial Framework uses the same broad principles set out in the Global Framework. Information below on Australia's Statistical Spatial Framework shows how these principles have been applied in the Australian context.

In addition, other countries such as Sweden, Mexico, New Zealand, Egypt and South Africa have produced similar frameworks for use nationally. These examples will be explored further below.

At a regional level, the GEOSTAT 3 project is looking at specifying and implementation of Global Statistical Geospatial Framework for the European Statistical System (ESS). The GEOSTAT 3 Project moves from generic principles to concrete recommendations, and with specific regional conditions to build on, including: INSPIRE, the ESS, the EFGS, EuroGeographics and UN-GGIM: Europe. The aim is to harmonize methods for the integration of statistical and geospatial information within the ESS; to modernize the statistical system and increase efficiency and flexibility in terms of output; and to provide a better foundation for collaboration between national statistical and geospatial agencies in providing society with more and better data for evidence based decision-making. The goal is for a fully geocoded round of population census, and provision of data for the SDG global indicator framework. More information on progress of the GEOSTAT 3 project can be found at the European Forum for Geography and Statistics website.[16]

Country example – Sweden

Statistics Sweden has written a planning report to assess the Global Statistical Geospatial Framework from a Swedish perspective.[17] Each of the five principles of the framework is assessed from a capability point of view. The assessment identified where there is room for improvements to be made and where there is a need for new tasks to be performed by Statistics Sweden. To understand the need for increased data integration the Global Framework is introduced as part of the solution for delivering new statistics as input for measuring the Sustainable Development Goals.

Integration of statistical and geospatial data has been done in Sweden since the 1980's, evolving over time and becoming more advanced as the tools and the skills allow for more geospatial analysis. Geospatial data, methods and processes are despite this not commonly known at Statistics Sweden and not well described in guiding documents.

In recent years Statistics Sweden has increased its international engagement, both in the UN Global Geospatial Information Management, UN-GGIM, and its global expert group on integration and through the coordination of EU-funded projects on integration of statistical and geospatial

[16] https://www.efgs.info/geostat/geostat-3/
[17] https://www.efgs.info/geostat/geostat-3/

information. The adoption of a Global Statistical Geospatial Framework gives the right platform to describe and improve the way Statistics Sweden work with statistical geospatial data integration.

The report includes a number of ideas for new tasks and for improvements, as well as statements of where it is good enough to maintain the current situation and where it could be possible to reduce efforts. These activities will be included in the yearly work plans for internal operations. Statistics Sweden is already collaborating with Lantmäteriet and other geospatial and earth observation data providers, but by implementing the GSGF they aim to will hopefully identify more areas for development and improvement.

Country example - Australia

The Australian Bureau of Statistics (ABS) recognised some time ago the need for and challenge of better integration of geospatial and statistical information and responded by developing the Statistical Spatial Framework (SSF). This Framework provides Australia with a common approach to connecting people-centric (socio-economic) information to a location, and improves the accessibility and usability of geospatially-enabled statistics.

What is the Statistical Spatial Framework?

The generic Statistical Spatial Framework, developed by the ABS, consists of five principles that are considered essential for integrating geospatial and socio-economic information (see green layers in diagram below). The Australian application of the Statistical Spatial Framework details the Australian implementation of the generic framework principles (see blue layers in Figure 6 below).

The Framework is also supported by international and Australian standards, and a range of guidance material; shown in Figure 7 on the following page.

Figure 6: Australian Statistical Spatial Framework

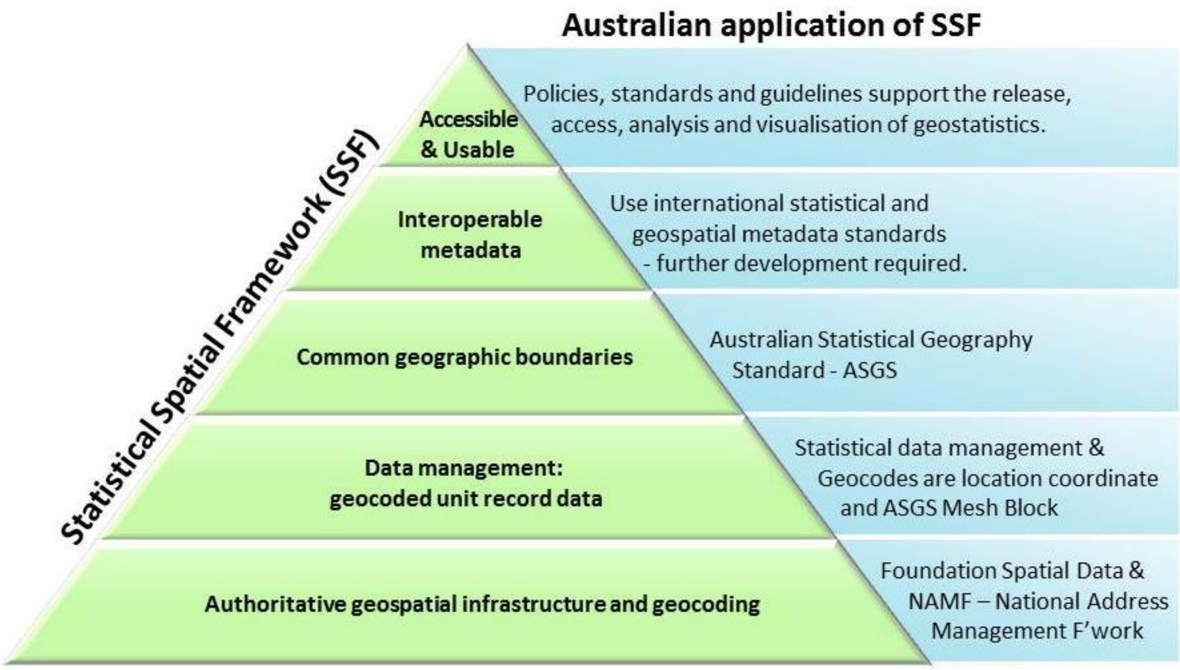


Figure 7: SSF on a page

Australian Bureau of Statistics

Statistical Spatial Framework

SSF Vision	Informed decision making is enhanced by using location in a common framework to allow seamless integration of administrative, statistical and spatial information resources.				
SSF Goals	All statistical data is consistently spatially enabled		Users can discover, access, integrate, analyse and visualise statistical information seamlessly for regions of interest		
SSF Principles	Authoritative geospatial infrastructure and geocoding	Data management – geocoded unit record data	Common geographic boundaries	Interoperable Metadata	Accessible and usable geostatistics
Australian Application of SSF	Use Foundation Spatial Data and relevant National Address Management Framework (NAMF) protocols	Geocodes stored on unit records are location coordinates and ASGS Mesh Blocks. Use statistical data management frameworks.	Data is released for Australian Statistical Geography Standard (ASGS) Statistical Area regions – as a minimum	Use international statistical and geospatial metadata standards[1]	Policies, standards and guidelines support the release, access, analysis and visualisation of spatially enabled information[1]
SSF Outcomes	• Consistent address and location information • Consistent geocoding results • Consistent management of geocoding issues	• Consistent and interpretable geocode information • Simplified aggregation of data to regions • Flexibility in production of regionalised data • Effective data management, especially privacy and metadata • Clear maintenance and custodianship	• Datasets can be integrated using common geography • Use of ASGS Statistical Areas simplifies visualisation and analysis • Metadata for ASGS and other regions supports data integration and use • Conversion of data between region types is supported	• Creation, discovery, integration and use of geostatistics is supported by statistical and geospatial metadata • Use of international standards enables the application of a larger pool of technologies, and wider data access and comparisons	• Data custodians can release data with confidence • Data users can discover and access geostatistics, and undertake analysis and visualisation • Web services enable machine to machine access and dynamic linking
Australian SSF - core standards, infrastructure and processes	• PSMA G-NAF, basemap and cadastre • NAMF & Addressing component of AS 4590:2006 • Geocoding guidance material • Point-of-entry address validation	• Unprojected coordinates (datum GDA94) • ASGS Mesh Blocks • Geocode metadata[1] • Mesh Block allocation tables	• ASGS - classifications and boundaries • Guidance material on use of regions • Metadata to support dissemination regions[1] • Geographic correspondences	• Statistical metadata frameworks[1] - SDMX and DDI • Geospatial metadata standard - ISO19115 • Semantic web (emerging)[1]	• Open data principles and policies • Web Services – OGC common standard and SDMX compliant • Guidance materials on - Privacy - Dissemination[1] - Visualisation - Analysis[1]

[1] Aspects of this component require further work.

Statistical Spatial Framework Principles – Goals, Objectives and Standards

The Statistical Spatial Framework closely aligns with the Global Statistical Geospatial Framework discussed in this paper. In particular, both frameworks share common goals, objectives and standards. To illustrate how the Global Framework is able to be adapted to different national contexts, the information below shows how Australia applies a range of national and international standards, frameworks, best practice and infrastructure for each principle.

Authoritative geospatial infrastructure and geocoding

- Authoritative geospatial data from the relevant National Spatial Data Infrastructure.

  The Foundation Spatial Data Framework (FSDF) is the authoritative source of geospatial data for Australia.

- Nationally consistent address collection standards to effectively capture the physical address.

  Australia has the National Address Management Framework (NAMF) that provides a coordinated approach to address management and geocoding, and uses the AS 4590 address standard.

- Point-of-entry address validation (for computer and internet based capture) will improve address quality and the resulting geocoding of information and also reduce the time spent correcting addresses after initial address data capture.

  Point-of-entry address validation is rapidly becoming accepted as best practice in Australia and around the world; and the SSF promotes this practice.

- Adopting common and practical methods of address data capture is seen as a key element of improving address data in administrative and collection based datasets.

  The Australian Bureau of Statistics (ABS) is seeking to have simplified and standardised data capture fields incorporated into addressing standards, and these currently form part of the SSF guidance materials.

- Common geocoding practice.

  The SSF promotes common geocoding practice through its guidance material.

Data Management

- Agreed statistical and geospatial data management frameworks.

  The ABS applies relevant National Statistical Organisation, as well as domestic and international, data management frameworks. The SSF promotes the use of the frameworks.

- Geocoding metadata standards.

  The SSF provides some guidance on the capture and storage of geocoding metadata; however, more comprehensive and internationally consistent standards are ultimately required.

- National privacy laws and agreed privacy standards.

  Australia has national privacy laws and principles that apply to the storage of data; the SSF promotes these laws and standards.

- Common geographic classifications.

  The ABS uses Mesh Blocks as a geocode type. Mesh Blocks are the smallest level of Geography in the Australian Statistical Geography Standard (ASGS).

- Geospatial data storage and national Geodetic Reference Frame.

  The current standard for the SSF is to store unprojected coordinates (datum GDA94).

Common Geography

- Common geographic classifications.

  The ASGS is Australia's statistical standard for the release of socio-economic data geospatially; the SSF encourages its use for the release of socio-economic data by all custodians.

- Accepted standards or guidance on the use of geographies for dissemination of data.

  The ABS provides guidance on the use of geographies through the ASGS and through guidance material that supports the SSF.

- International statistical and geospatial metadata standards.

  The ABS applies relevant statistical and geospatial metadata standards, and the SSF promotes the use of these standards.

- National privacy laws and agreed privacy standards.

  Australia has national privacy laws and principles that apply to the release of data; the SSF promotes these laws and standards.

- Systems and methods to correspond data between geographies.

  The ABS produces correspondences allowing data to be converted between ASGS geographies and other geographies.

The SSF promotes the following ideal geographic classification principles:

- *Has a hierarchical design.*
- *Supports flexible reporting by including a building block of small geographic areas, such as a suburban block, suburb or neighbourhood; which can then be aggregated to larger areas, such as natural resource management areas or larger administrative geographies.*
- *Is designed to include a hierarchy of geography types that contain approximately equal population numbers within each area of a geography type.*
- *Is stable over time.*

Statistical and Geospatial Metadata Interoperability

- International statistical and geospatial metadata standards.

  The ABS applies relevant statistical and geospatial metadata standards; however, there is currently an acknowledged gap in interoperability between these standards that needs to be addressed internationally. The SSF promotes the use of the existing standards.

Accessible and usable Geostatistics

- International principles or agreed national policies on open data.

  The ABS applies relevant national open data policies and principles, and the SSF also promotes the use of these policies and principles.

- Accepted standards or guidance on the use of *geographies* for dissemination, visualisation and analysis.

  The ABS and the SSF promote the use of a range of resources for dissemination, visualisation and analysis of data, including UN resources.

- Web service standards.

  Open statistical/geospatial standards developed by the Open Geospatial Consortium (OGC) and Statistical Data and Metadata Exchange (SDMX) are promoted through the SSF.

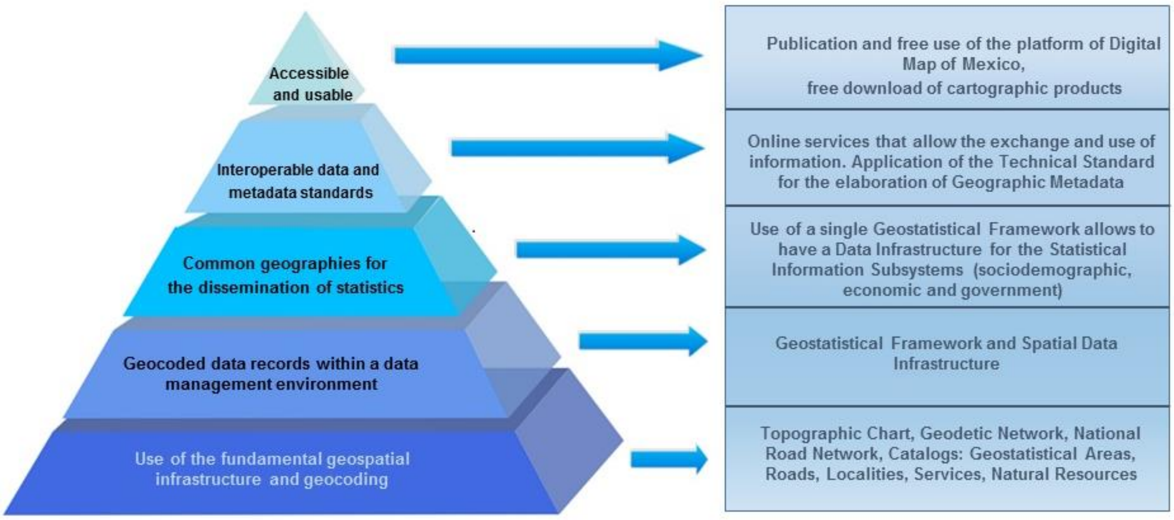
- National privacy laws and agreed privacy standards.

  Australia has national privacy laws and principles that apply to the access and use of data; the SSF promotes these laws and standards.

Country example - Mexico

Mexico's National Geostatistical Framework is an example of a national system that geocodes statistical information from their censuses and surveys to a location and different levels of geography.

The National Geostatistical Framework aligns with the Global Statistical Geospatial Framework discussed in this paper. Following the example of the Australian application of the Statistical Spatial Framework, the information below shows how Mexico applies a range of national and international standards, frameworks, best practice and infrastructure for each principle.



Use of Fundamental geospatial infrastructure and geocoding

- Authoritative geospatial data from the relevant National Spatial Data Infrastructure.

  By law National Institute of Statistic and Geography (INEGI by its name in Spanish, Instituto Nacional de Estadística y Geografía) is the authoritative source of geospatial data used in Mexico for georeferencing statistical information.

- Nationally consistent address collection standards to effectively capture the physical address.

  Mexico has a Technical Standard to describe the home, based on the Geostatistical Framework for Mexico or the topographic map.

  In addition, there are catalogues and cartography that supports the process of georeferencing.

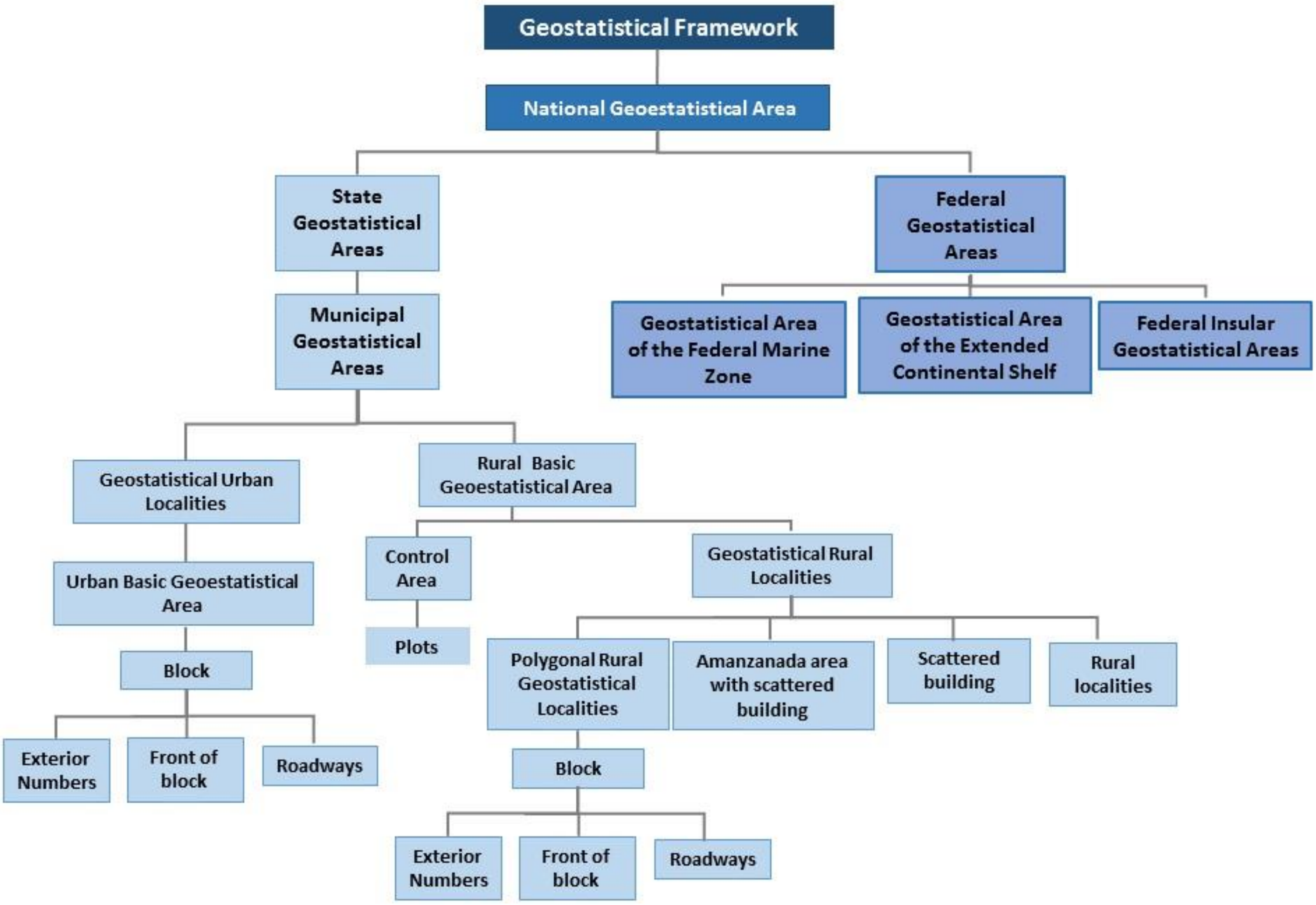
- Adopting common and practical methods of address data capture is seen as a key element of improving address data in administrative and collection based datasets.

INEGI works with the various levels of government to advise on the adoption of the Technical Standard for Geographic Address for Mexico.

INEGI is also developing systems to approach the correct adoption of the Technical Standard.

- Common geocoding practice.

INEGI enables different levels of government to adopt the Technical Standard for Geographic Address.



Geocoded unit record data in a data management environment

- Agreed statistical and geospatial data management frameworks.

The Geostatistical Framework has been adopted across different levels of government for various administrative and statistical data analysis applications.

- Geocoding metadata standards.

The Geostatistical Framework is composed of catalogs for geocoding statistical data. Similar to the standards reported by Australia, Mexico seeks consistency with international standards.

- National privacy laws and agreed privacy standards.

There are a number of current laws that apply to personal information and INEGI has published several principles for statistical information.

- Common geographic classifications.

According to the principles of confidentiality of statistical information, the most disaggregated level of information is the Block (for urban areas) and the Production Unit (for rural areas).

- Geospatial data storage and national Geodetic Reference Frame.

The Geostatistical Framework is stored in Lambert Conformal Conic projection, using as a reference frame ITRF 2010 epoch 2010.0.

Common geographies for dissemination of statistics

- Common geographic classifications.

Currently in Mexico, all statistical information is linked to the Geostatistical Framework and it is based on the Geodetic Reference Framework that is used for all geospatial and environmental information, making it possible to link the statistics and geographical variables.

- Accepted standards or guidance on the use of geographies for dissemination of data.

INEGI has standards and guidelines for the publication and dissemination of statistical and geographical information.

- International statistical and geospatial metadata standards.

INEGI uses standards to prepare statistical and geographical metadata; these standards are used by other producers of statistics and geographic information. Particularly, the standards refer the FGDC-STD001-1998 and ISO 19115:2003

- National privacy laws and agreed privacy standards.

Mexico has federal laws for confidential data and INEGI has published standards for statistical and geospatial data.

Interoperable data and metadata standards

- International statistical and geospatial metadata standards.

  INEGI applies standards for statistical and geospatial data. INEGI agrees with Australia on the need to seek international agreement.

Accessible and usable geospatially enabled statistics

- International principles or agreed national policies on open data.

  INEGI applies principles and policies of open data and standards for the national interest.

- Accepted standards or guidance on the use of *geographies* for dissemination, visualisation and analysis.

  INEGI promotes mechanisms for dissemination, visualization and data analysis.

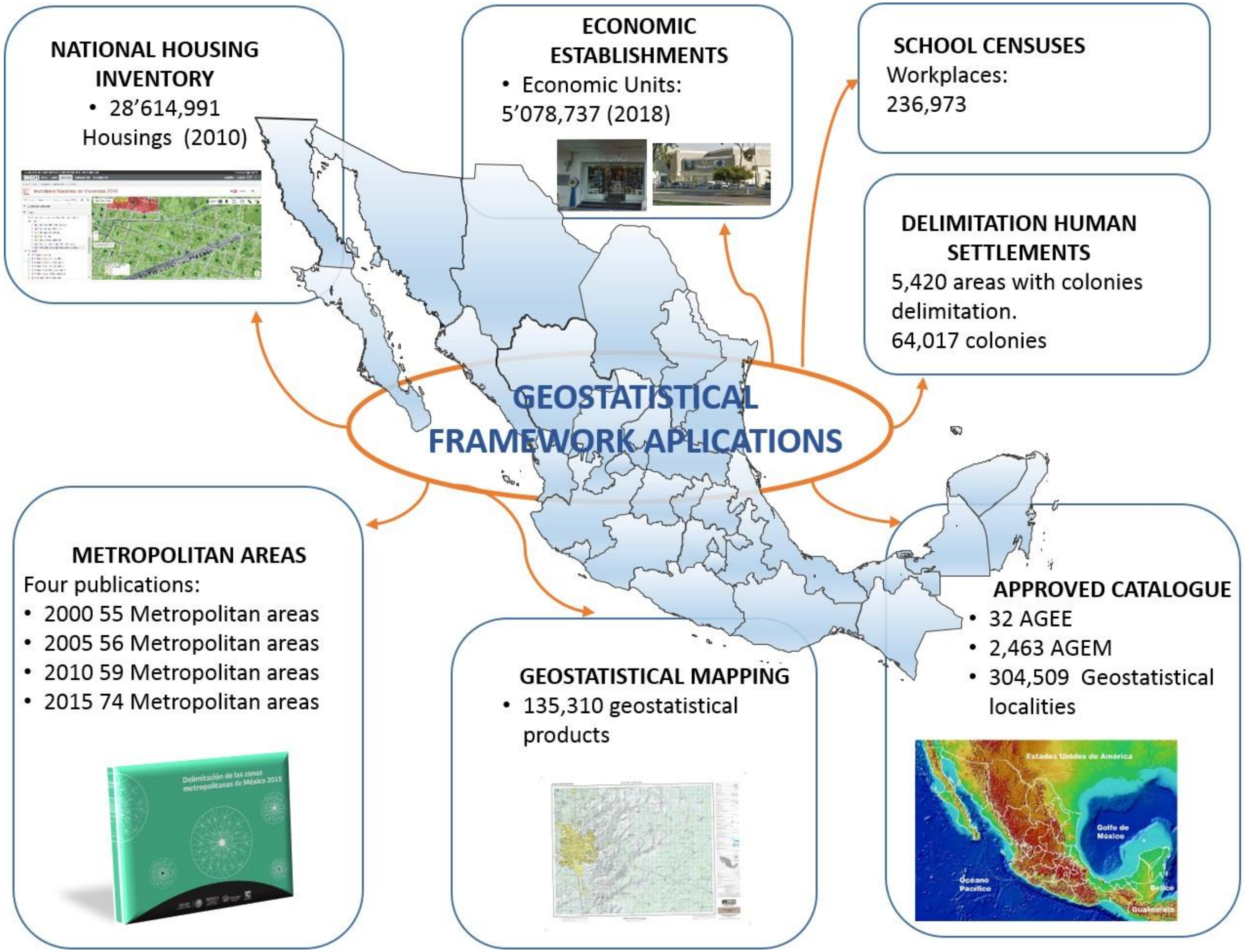
- Web service standards.

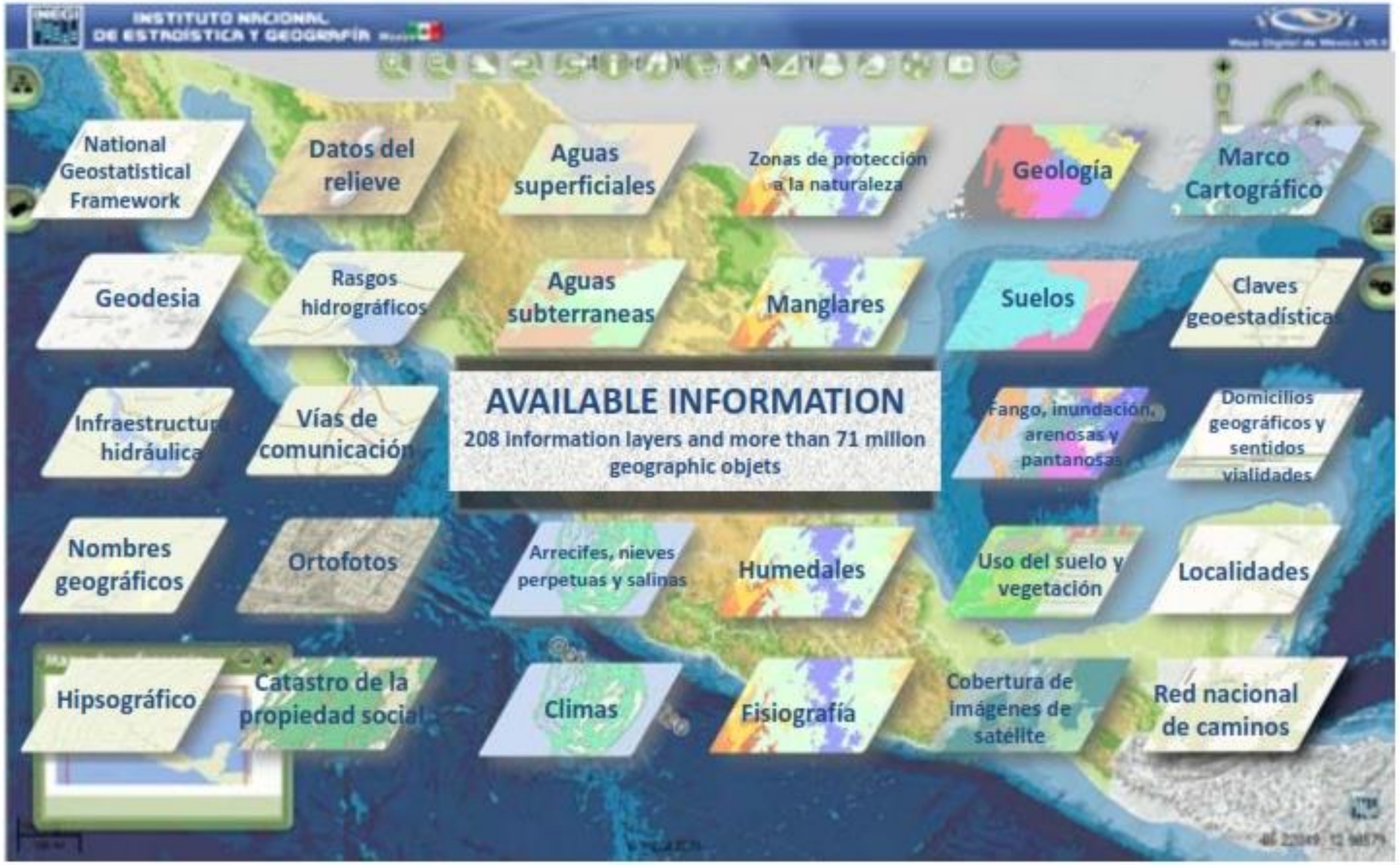
  INEGI promotes the adoption of open data based on the recommendations from OGC and SDMX.

- National privacy laws and agreed privacy principals.

  Mexico has laws on information security and INEGI has published the laws related to the statistical and geographical data.

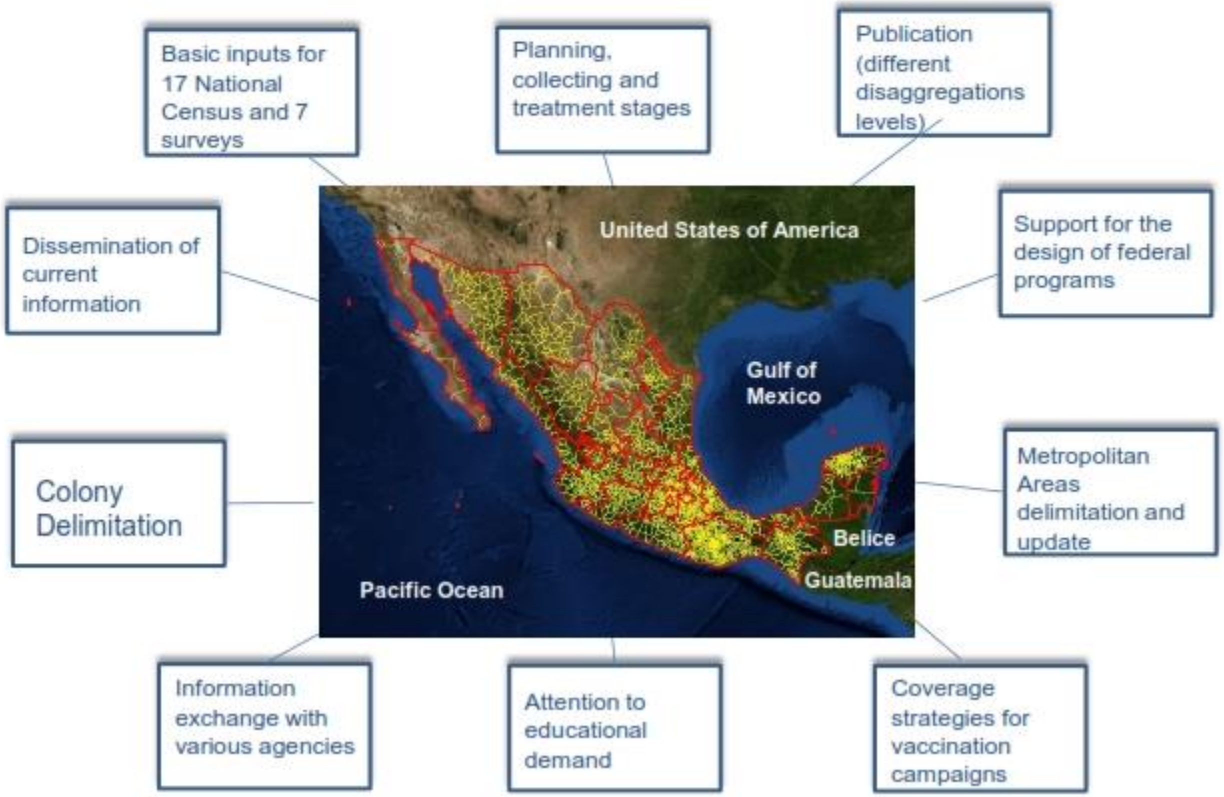
The diagrams on the following pages show how Mexico uses their National Geostatistical Framework.


NATIONAL HOUSING INVENTORY
• 28’614,991 Housings (2010)
ECONOMIC ESTABLISHMENTS
• Economic Units: 5’078,737 (2018)
SCHOOL CENSUSES
Workplaces: 236,973
DELIMITATION HUMAN SETTLEMENTS
5,420 areas with colonies delimitation.
64,017 colonies
GEOSTATISTICAL FRAMEWORK APLICATIONS
METROPOLITAN AREAS
Four publications:
• 2000 55 Metropolitan areas
• 2005 56 Metropolitan areas
• 2010 59 Metropolitan areas
• 2015 74 Metropolitan areas
Delimitación de las zonas metropolitanas de México 2015
GEOSTATISTICAL MAPPING
• 135,310 geostatistical products
APPROVED CATALOGUE
• 32 AGEE
• 2,463 AGEM
• 304,509 Geostatistical localities
Estados Unidos de América
Golfo de México
Océano Pacífico
Belice
Guatemala



INEGI
INSTITUTO NACIONAL DE ESTADÍSTICA Y GEOGRAFÍA

NATIONAL GEOSTATISTICAL FRAMEWORK USES

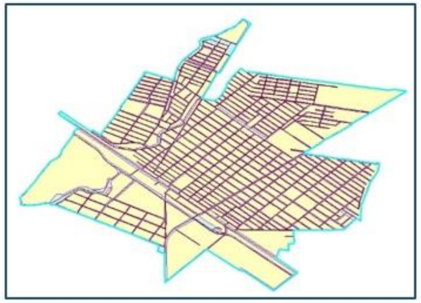


INEGI
INSTITUTO NACIONAL DE ESTADÍSTICA Y GEOGRAFÍA

NATIONAL GEOSTATISTICAL FRAMEWORK DESEGREGATION CURRENT LEVELS

32 State Geostatistical Areas (AGEE)

2,463 Municipal Geostatistical Areas (AGEM)



4,563 Geostatistical Urban Localities

49,720 Polygonal Geostatistical Localities

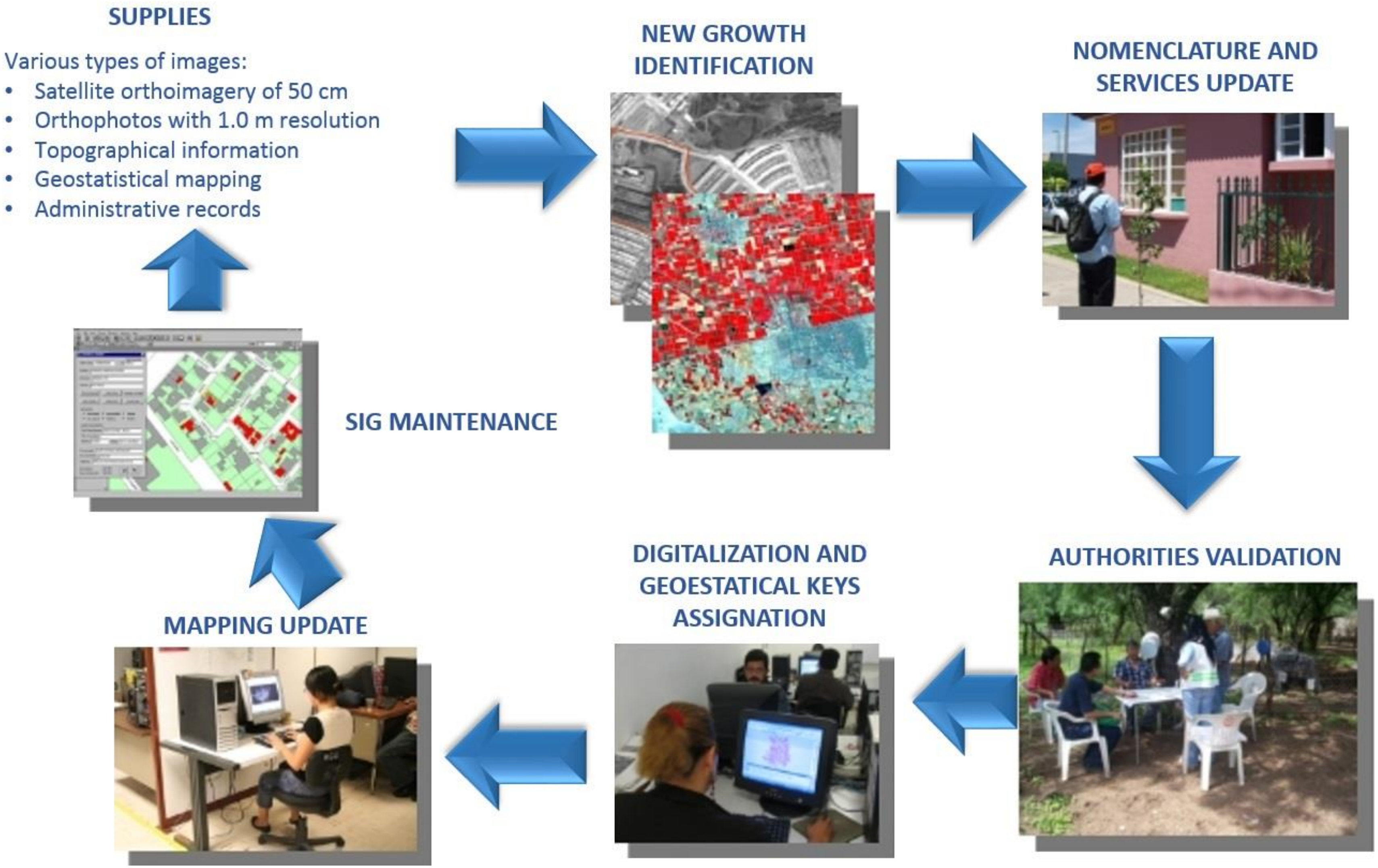
10’922,312 Roadways

1’481,763 Urban localities blocks

868,089 Rural localities blocks

38’010,866 Total exterior numbers

PERMANENT UPDATING FRAMEWORK PROCESS



Country example - New Zealand

Statistics New Zealand's new Statistical Units Model is an example of a national system that geocodes statistical information from their censuses and surveys and produces statistics for different levels within a geography standard.

The Statistical Units Model aligns with the Global Statistical Geospatial Framework discussed in this paper. Following the examples from Australia and Mexico, the information below shows how New Zealand applies a range of national and international standards, frameworks, best practice and infrastructure for each principle.

Use of fundamental geospatial infrastructure and geocoding

- Authoritative geospatial data from the relevant National Spatial Data Infrastructure.

  The digital cadastral database is the authoritative source of geospatial data for New Zealand. Legislation governs higher-level geographies for local governments and electoral purposes.

- Nationally consistent address collection standards to effectively capture the physical address.

  Different addressing standards exist to meet differing business needs. New Zealand uses the AS/NZS 4819:2011 Rural and Urban addressing standard for assigning addresses and ISO 19160-1:2015 Addressing-Part 1: Conceptual Model which provides a conceptual foundation for address data.

- Point-of-entry address validation (for computer and internet based capture) will improve address quality and the resulting geocoding of information and also reduce the time spent correcting addresses after initial address data capture.

  Point-of-entry address validation is desired but currently no address databases validate address at point-of-entry to an authoritative source.

- Adopting common and practical methods of address data capture is seen as a key element of improving address data in administrative and collection based datasets.

  A national address exchange standard has been developed. Statistics NZ and Land Information NZ have implemented a New Zealand Profile of ISO 19160, an addressing exchange standard.

- Common geocoding practice

  Statistics NZ have created a statistical location register and the supporting infrastructure to enable common geocoding practice across the organisation. The corporate geocoder enables datasets to be matched against an address reference source and obtain a persistent id. This enables the ability to better integrate data and compare and do analysis.

Geocoded unit record data in a data management environment

- Agreed statistical and geospatial data management frameworks.

  Statistics NZ applies relevant national statistical data management frameworks, as well as domestic and international data management frameworks.

- Geocoding metadata standards.

  Statistics NZ has not implemented any geocoding metadata standards

- National privacy laws and agreed privacy principles.

  The Statistics Act 1975 governs Statistics NZ's approach to maintaining secrecy. The Privacy Act (1993) underpins a lot of what we do at Statistics NZ, including how we collect data, and how we protect the personal information of our own people.

- Common geographic classifications.

  Statistics NZ currently uses the Meshblock as its common geographic classification. Meshblocks are the smallest statistical spatial unit for statistical, electoral and government administration in New Zealand.

- Geospatial data storage and national Geodetic Reference Frame.

  Data is stored centrally in the New Zealand Transverse Mercator projection using the NZ geodetic datum 2000

Common geographies for dissemination of statistics

- Common geographic classifications.

  Statistical standard for geographic areas 2018 (SSGA18) is Stats NZ's official standard for statistical geographic areas in New Zealand. It allows statistical units, such as households, people, or businesses, to be assigned to the location where they live, work, and operate.

  SSGA18 enables the production of integrated statistics by geographic area. It provides a range of geographic units that are convenient for data collection, compilation, and output, and are useful for spatial analysis of social, demographic, and economic statistics.

- Accepted standards or guidance on the use of *geographies* for dissemination of data.

  The SSGA18 is a hierarchical geography that ranges from Meshblock, SA1, SA2, Territorial Authority, Regional Council. Although the lowest level is the meshblock, the new SA1 output geography will allow the release of more low-level data than is available at the meshblock level.

In the urban rural geography, urban areas have been redesigned to represent the urban 'footprint', i.e. areas of high population density. Rural areas are broken down by rural settlement (previously rural centres in the census area unit geography) and other rural.

- International statistical and geospatial metadata standards.

  Statistics NZ uses AS/NZS ISO 19115:2005, Geographic information — Metadata to document geographic boundaries.

- National privacy laws and agreed privacy principles.

  The Statistics Act 1975 governs Statistics NZ's approach to maintaining secrecy. The Privacy Act (1993) underpins a lot of what we do at Statistics NZ, including how we collect data, and how we protect the personal information of our own people.

- Systems and methods to correspond data between geographies.

  Statistics NZ produces concordances allowing data to be converted between annual Meshblock versions and all statistical, electoral and government administration areas.

Interoperable data and metadata standards

- International statistical and geospatial metadata standards.

  Statistics NZ applies relevant statistical and geospatial metadata standards; however, there is currently an acknowledged gap in interoperability between these standards that needs to be addressed internationally.

Accessible and usable geospatially enabled statistics

- International principles or agreed national policies on open data.

  Statistics NZ complies with the New Zealand Government Open Access and Licencing (NZGOAL) framework for data releases.
  New Zealand signed up to the International Open Data Charter in 2017. The charter supports and builds on the New Zealand Declaration on Open and Transparent Government and the supporting Data and Information Management Principles.

  The charter has six principles and supporting actions. Data should be:

  - *open by default*
  - *timely and comprehensive*
  - *accessible and usable*
  - *comparable and interoperable*
  - *for improved governance and citizen engagement*
  - *for inclusive development and innovation.*

- Accepted standards or guidance on the use of *geographies* for dissemination, visualisation and analysis.

  Statistics NZ releases annual geographic classification, concordances and boundaries for use in dissemination, visualisation and analysis.

- Web service standards.

  Statistics NZ are providing open geospatial web services for some products as a service.

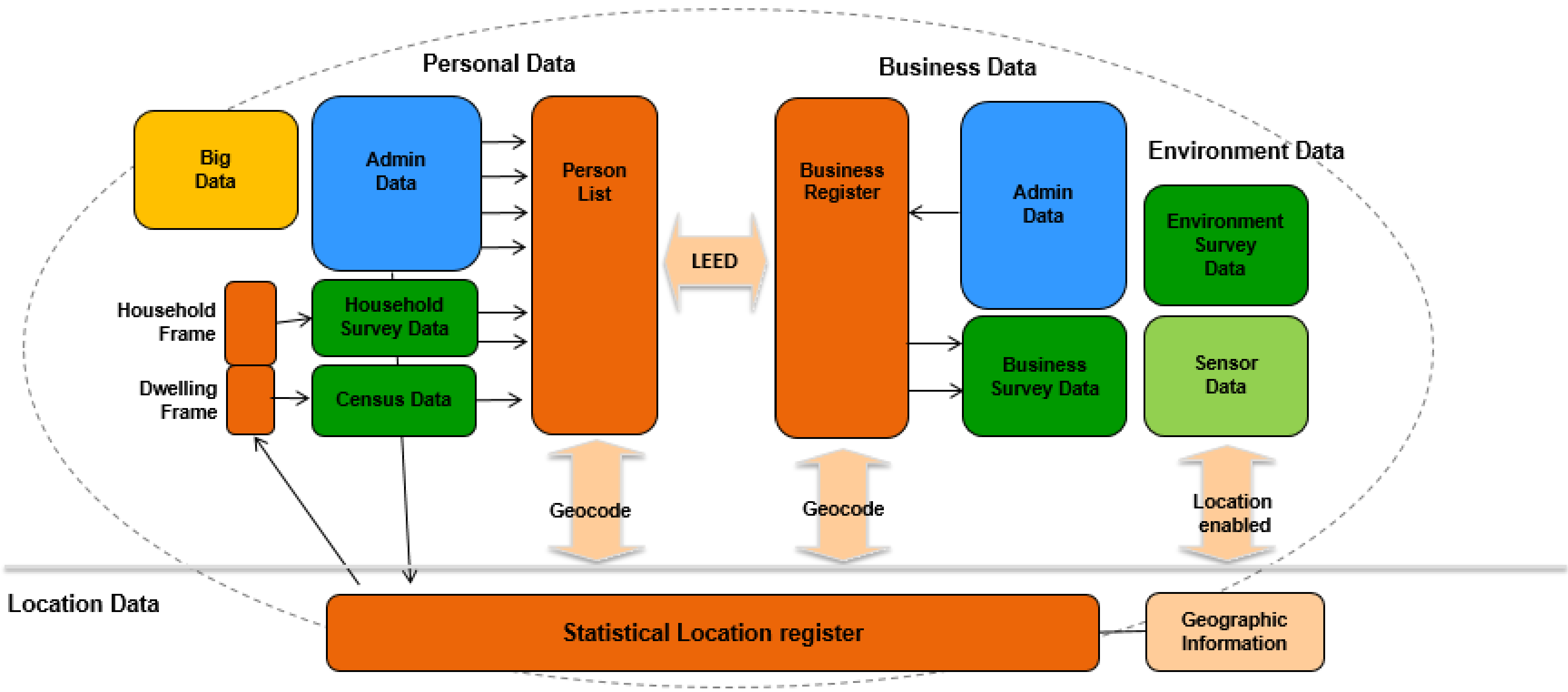
- National privacy laws and agreed privacy principles.

  The Statistics Act 1975 governs Statistics NZ's approach to maintaining secrecy.

The diagrams on the following pages show how Statistics New Zealand uses their Statistical Units Model.

Potential Future Statistical Units Model

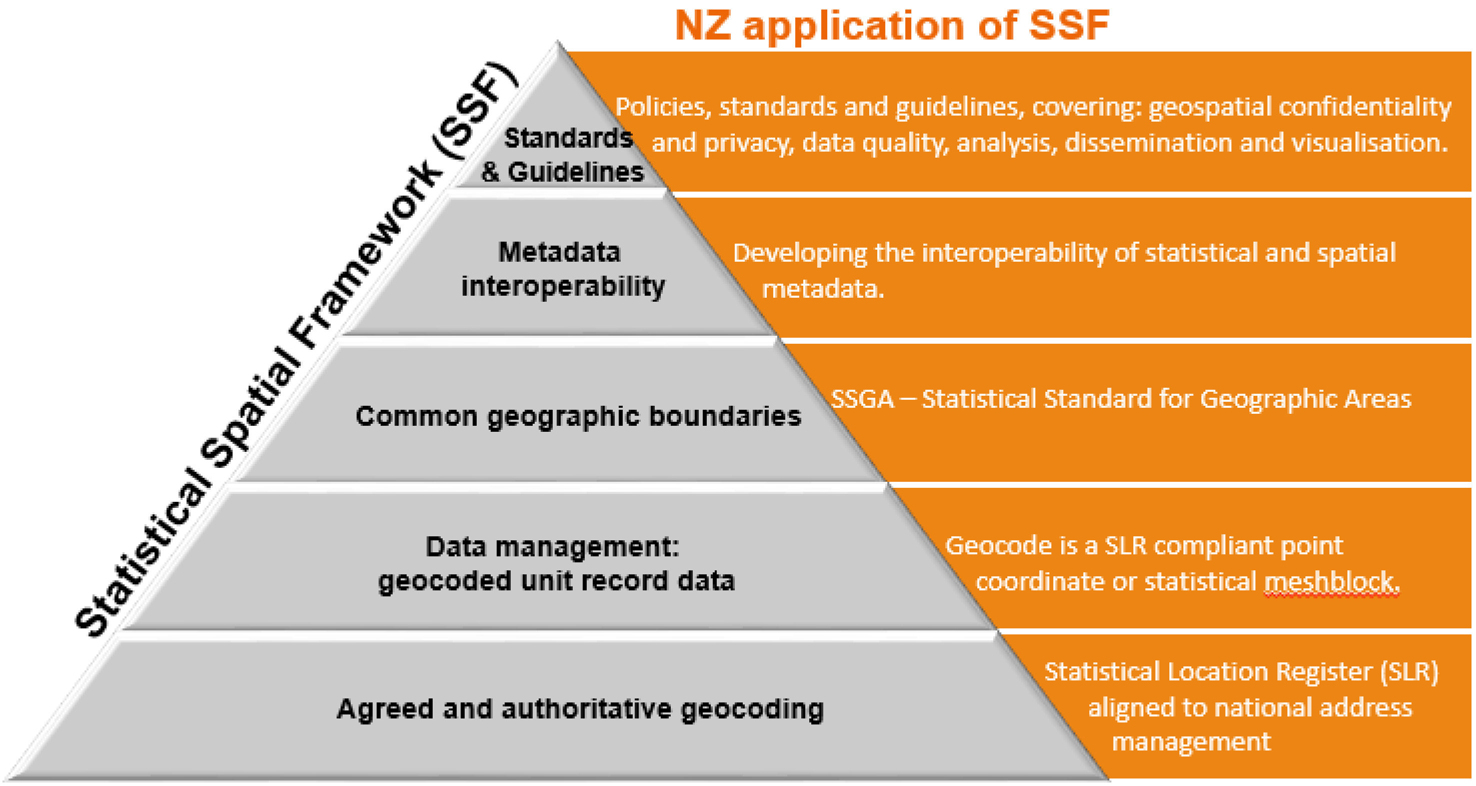
Geo-enabled statistical units model – linking data to and by location



Statistical Spatial Framework (SSF)

Stats NZ
Tatauranga Aotearoa

Integration of statistical & geospatial information



Country example – Egypt

Use of fundamental geospatial infrastructure and geocoding

- Authoritative geospatial data from the relevant National Spatial Data Infrastructure (NSDI).

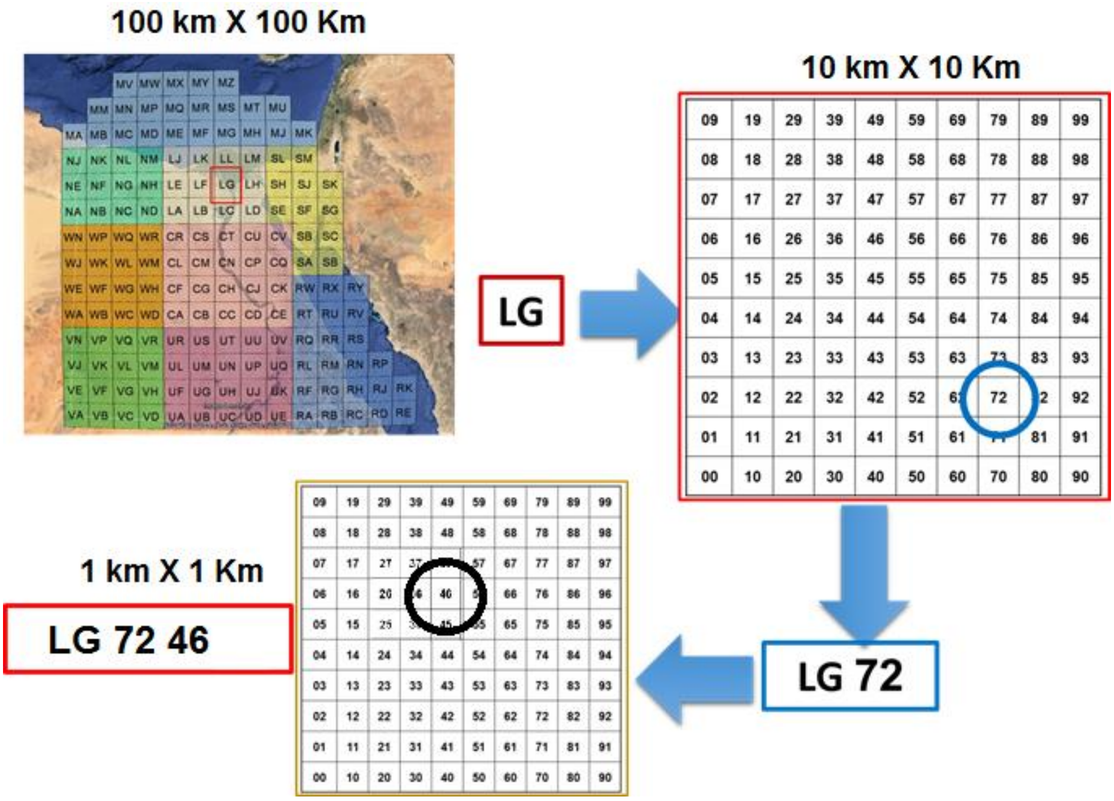
Egypt is working on establishing a National Spatial Data Infrastructure using a National Grid system with coordinate system MTM-WGS84 (Modified Transverse Mercator), which will be used to generate a unique numbering and geocoding system for each unit in a dataset, such as a building, household or business.

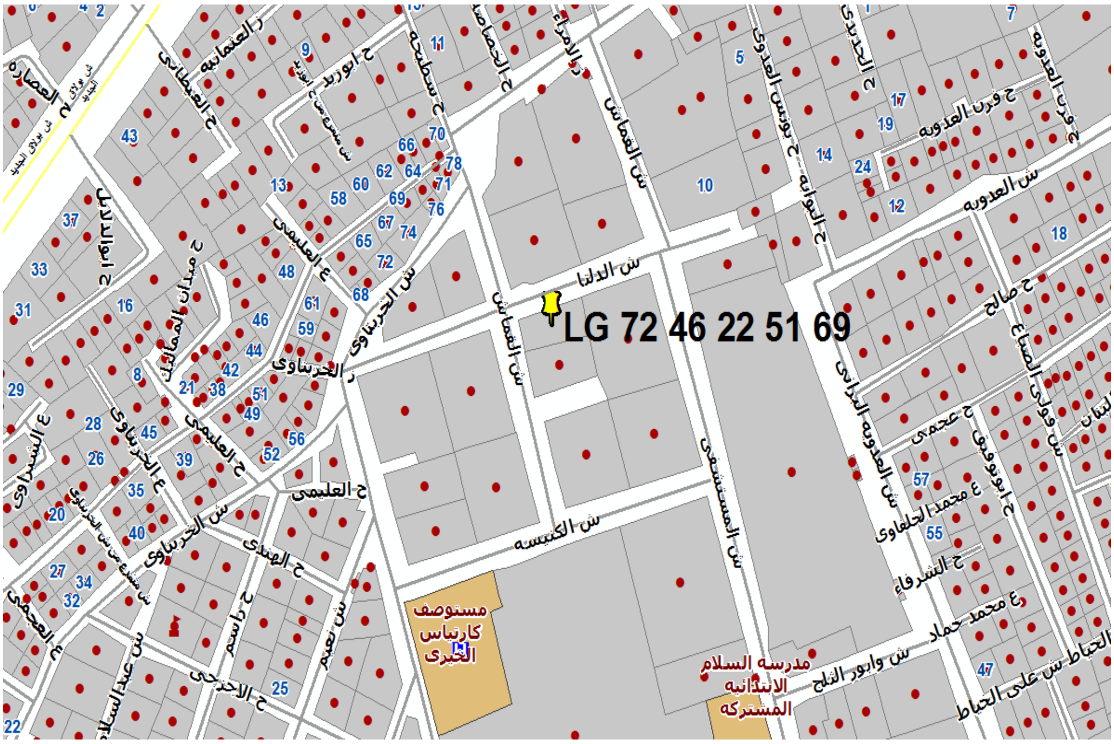
How this geocoding system is generated?

First a grid based system was generated to cover the Egyptian land in different scale levels

*First level is a grid each cell is 100 km * 100 km and by taking any cell of them for example LG cell it is divided into cells each one of them is 10 km * 10 km and by taking any cell of them for example 72 this cell name is LG 72 and this cell it is divided into cells each one of them is 1km * 1 km and by taking any cell of them for example 46 this cell name is LG 72 46 and So on till reaching 1 m * 1m .*

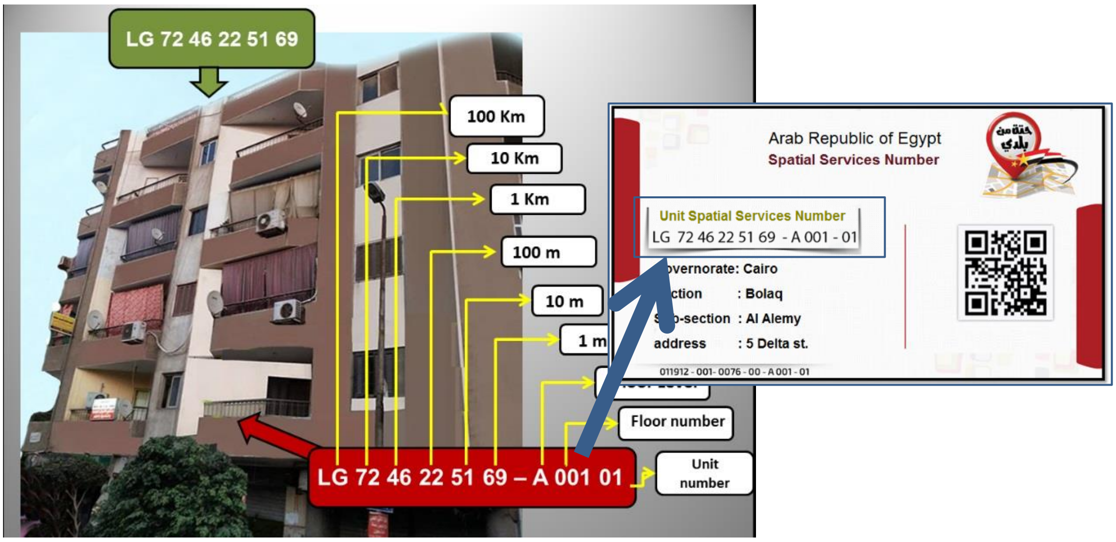
Grid Based System





This method is applied automatically on the digital maps to produce a unique numbering for each building and unit linked to its geographical location which does not change with the administrative boundaries change.

This number is named as "Spatial ID" and distributed on individuals during the enumeration stage of Census 2017



Geocoded unit record data in a data management environment

- Geospatial data storage and national Geodetic Reference Frame.

All the geographical spatial data and statistical data in Egypt are identified with the smallest geographic boundary (Shyakha (Urban) - Village (Rural)). Egypt is currently working now on putting a mechanism in place to facilitate the integration and management of the geocode within a dataset including spatial number identifier to geocode linking mechanisms. This will join the detailed statistical building, household and business data. This will be accomplished by the end of the next census in 2016.

National Spatial Identifier is considered a new Geocoding System

- *Is a unique number for each unit within the building.*
- *Is produced automatically from digital maps.*
- *Giving a fixed number for the unit linked to its geographical location which does not change with the administrative boundaries change.*

After finishing the Census it was a great opportunity to integrate census data with the Egyptian NSDI, as Egypt is working now on establishing a NSDI

The only way to make this integration is using the Spatial Id as a common unique key linking all national databases of the country to organize, evaluate and follow-up of local government services such as electricity, natural gas, and sewage.

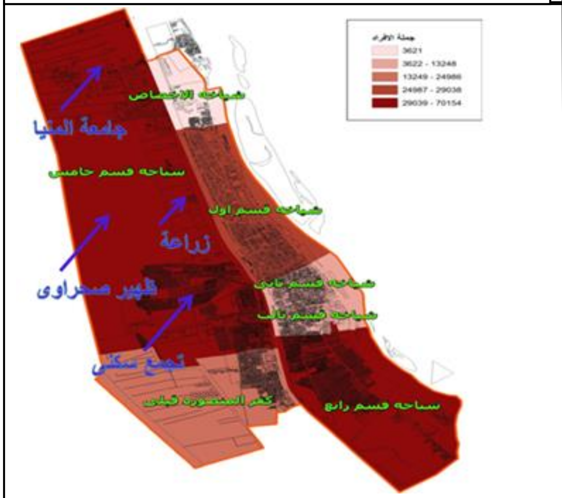
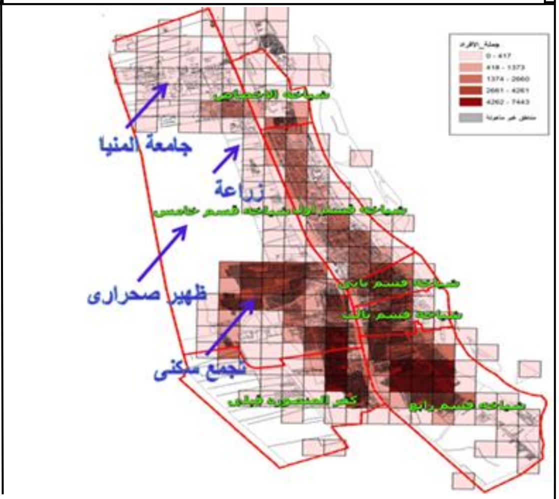
Common geographies for dissemination of statistics

- Common geographic classifications.

One of the most important roles of the National Spatial Data Infrastructure is to unify the administrative geographic boundaries between all the Egyptian agencies. It is also used as a mechanism for data dissemination using the National grid merged with the geographic boundaries.

<u>Case Study: National grid merged with the geographic boundaries and</u>

<u>The comparison between geospatial data dissemination with using administrative boundaries and national grid system</u>

Total Population at the level of Shiakha of Minea City	*Total population Grid Square 250*250 Of Minea City*
	

<u>The advantages of using national grid system:</u>

- *It does not depend on administrative or topographical boundaries.*
- *It Fixed and regular areas.*
- *Ease of data integration for different grid squares.*
- *It distinguishes between urban and non-urban areas.*
- *Realistic and accurate results to support decision makers.*

Interoperable data and metadata standards

- International statistical and geospatial metadata standards.

Egypt applies standards for statistical and geospatial data, and agrees with Australia on the need to seek international agreement.

One of the main targets of Egypt NSDI is to create a standard unified base map for Egypt to be used by all the governmental authorities and now the Military Survey Authority is working on building this base map using a recent ortho photo with a high resolution and applying the Spatial ID which produced by CAPMAS to integrate with the census data.

Also NSDI community approved the maps of the administrative boundaries of Egypt witch modified and updated during Census2017 by CAPMAS to be the standard basemap of the administrative boundaries to use by all the governmental authorities.

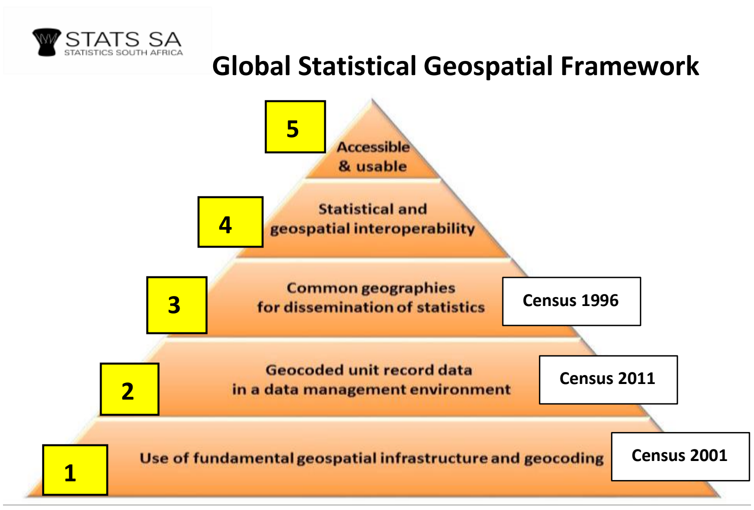
Accessible and usable geospatially enabled statistics

*After finishing the census CAPMAS starts the dissemination stage using the Egyptian geoportal and Egyptian development atlas by disseminating the results on the Egyptian Grid system in different results (100KM*100Km, 10KM*10Km, 1KM*1Km, and 200m*200m).*

Which gives high accuracy and precision for data analysis.

Country example – South Africa

South Africa's inspiration to integrate statistics and geo-spatial data began many decades ago; where the main purpose was to automate the mapping office that produced paper maps for the population censuses. This humble thinking has unknowingly resulted in South Africa, through the different population censuses, to progressively interact, with the ***Global Statistical Geospatial Framework (GSGF)***, as it is known today.



Immediately after Census 1996, joint projects with the National Mapping Agency (NMA) to create a cadastral base map, to digitise enumeration areas (Principle 1). Dissemination for Census 1996 made use of administrative hierarchy, first place-name/ community data dissemination was achieved (Principle 3). Use of SuperCross and map visualisation (Principle 5).

Continued collaboration to use fundamental geospatial data to demarcate enumeration areas for Census 2001 (Principal 1). Geocoded enumeration areas (Principle 2). Followed by Principle 3 and 5.

Geocoding of dwellings, dot on map linked to a record on a list (Principle 2) for Census 2011. Followed by Principle 3 and 5.

Since 2005, Stats SA has invested in the creation of a national dwelling frame. The dwelling frame is currently the centre of digital capturing for data collection; the household sampling frame, and the base collection frame for population censuses; including the upcoming 2021 census. In other words, it drives the organisation's modernisation thinking. Similar thinking goes around spatializing the national business register and the national population register.

The current data demands, driven mainly by the global, regional and local policy frameworks, requires this integration or embedding; provides the required evidence-base, and is the catalysts for innovative methods to fill data gaps. South Africa's 2030-National Development Plan recognises that overcoming the country's triple challenge of poverty, inequality and unemployment lies in transforming its physical space; national *spatial* development, strategies for *spatial* change, transforming *spatial* patterns and localising. Therefore, space emerges as a fundamental growth and development thrust for the country. Various frameworks and policies are being developed; and there is extreme need for (disaggregated) statistical and geospatial data.

The NSO and the NMA continue to innovate within tight budget constraints. The GSGF provides a guiding framework for consistent integration; and holistic thinking from disparate data sources to geospatially enabled statistics for evidence-based decision-making.

Use of Fundamental geospatial infrastructure and geocoding

- Authoritative geospatial data from the relevant National Spatial Data Infrastructure.

  South Africa has a Spatial Data Infrastructure Act since 2003 focussing on the establishment of the South African Spatial Data Infrastructure (SASDI). The Department of Rural Development and Land Reform (DRDLR) is mandated to administer the Act in collaboration with other stakeholders. The establishment of a Committee for Spatial Information is mandatory. Stats SA is a compulsory member. The Committee is functional with established Sub-Committees (Policy and Legislation, Education and Training, Standards, Systems, Marketing and Communication and Data).

  The Data Sub-Committee has so far drafted a list of base (fundamental) geospatial datasets for South Africa. The process to assign coordinating data custodians (bringing together common data themes to work together) and data custodians (in the different themes) has begun. As a start 10 base dataset themes are prioritised (administrative boundaries, imagery, transport, social statistics, land use, land cover, cadastral, hydrology, geodesy and conservation).

  The minimum list of Global Fundamental Geospatial Data Themes, as proposed by the UN-GGIM Working Group on Fundamental Geospatial Data Themes, is being used for alignment.

- Nationally consistent address collection standards to effectively capture the physical address.

  A South African National Standard for addresses was published around 2009. The Standard covers address formats, aspects of data interoperability, address allocations and updating.

  An address is an important service delivery point (utilities, emergency response, etc.), a requirement for opening bank accounts, for obtaining identity documents, for voting, for employment, etc. In South Africa there are several role players who allocate addresses (like the local municipalities and the South African Post Office). Stats SA also allocated temporary addresses in certain parts of the country during the censuses. The Standard aims to make all the different role player's data interoperable through standardized geocoding (and to reduce duplication and costs), and interoperable with other base data like the link with the national base cadastre.

  The Standard covers the different kinds of addresses for the different settlement types in South Africa, stemming from the country's history of compromised land use patterns (formal areas, informal areas, farm areas and traditional areas). It should be noted that valid or legitimate addresses are still to be assigned in parts of the country.

- Adopting common and practical methods of address data capture is seen as a key element of improving address data in administrative and collection based datasets.

  Adopting the Standard still remains a challenge in South Africa. Stats SA, for purposes of conducting censuses and for survey sampling requires a national database of addresses or dwellings. Address data is received from various stakeholders and compiled into a national database. In areas with no addresses, other geospatial data like imagery, as well as physical ground visits, are used to compile dwelling locations with their unit counts and dwelling type classification (like residential or commercial). In this way Stats SA has compiled and maintained over these years a national source of addresses or dwellings.

- Common geocoding practice.

  Through the SDI Act, the process of assigning a Coordinating Custodian (to pull all parties together) and formalizing the Custodians, together with approved policies (like the Data Custodianship Policy that clarifies amongst other aspects such as roles and responsibilities, frequency of updating, and data quality), it is envisaged that the data becomes authoritative, and can be maintained in a standardized manner.

Geocoded unit record data in a data management environment

- Agreed statistical and geospatial data management frameworks.

  Geospatial data management framework used by NMA – Geospatial data is collected, processed and managed using generally accepted international best practice and application of various OGC/ ISO Standards, such as ISO 19157 (Data Quality), ISO 19111 (Spatial Referencing), ISO 19115 (Metadata) and ISO 19144 (Classification systems). The national spatial reference system is based on the International Terrestrial Reference Frame (ITRF).

  Statistical data management framework used by Stats SA – Statistical Value Chain (SVC) approach (a modified version of the Generic Statistical Information Model (GSIM); it also has main components of the UNECE's Generic Statistical Business Process Model – needs, design, build, collect, process, analyze, disseminate, evaluate). The statistical modernization or digitalizing process planned for Stats SA must embed the geospatial thinking for all data collected.

- Geocoding metadata standards.

  ISO 19115(Metadata) is a core standard used.

- National privacy laws and agreed privacy standards.

  Stats SA adheres to the UN Fundamental Principles for official statistics (Principle 6).

  The Statistics Act clearly stipulates matters of data confidentiality and disclosure.

  In South Africa there is also the Protection of Personal Information Act (2013) (POPI).

The data release for Census 2011 did not include data at enumeration area level. Stats SA created and disseminated data for small-areas (a combination of enumeration areas to prevent data disclosure).

- Common geographic classifications.

Although enumeration area data is not released, the enumeration area layer is aggregated to various geographic classifications. Common aggregations are the administrative boundaries of the country; community/ places (like city/town/suburbs/village), urban, farm, traditional – urban/ rural; and other service delivery areas like health, policing, and education. Stats SA provides the standard links to these aggregations.

Not possible to demarcate and geocode enumeration areas (polygons) to too many administrative areas. The administrative boundaries were accepted as important, and enumeration areas geometrically align, and are geocoded to these boundaries. Other boundaries are usually approximated (point-in-polygon or majority rule) but the standard concordance files are supplied by Stats SA. It is envisaged that the point based system of data collection will solve this problem.

- Geospatial data storage and national Geodetic Reference Frame.

Geospatial data is stored in either vector data format with full topological structuring, gridded data or raster (imagery) data. The national geodetic reference frame provides for a unique spatial reference frame across the whole country. It is tied to the ITRF.

Common geographies for dissemination of statistics

- Common geographic classifications.

(See above)

- Accepted standards or guidance on the use of geographies for dissemination of data.

Stats SA provides the necessary concordance files to any aggregation developed by the organisation.

- International statistical and geospatial metadata standards.

Statistical metadata standards used by Stats SA - Statistical Data and Metadata Exchange (SDMX), and Data Documentation Initiative (DDI) for time series data released on StatsOnline.

Geospatial metadata standards used – (See above)

- National privacy laws and agreed privacy standards.

(See above)

Interoperable data and metadata standards

- International statistical and geospatial metadata standards.

  Statistical metadata standards used by Stats SA – In addition parts of ISO19115 are being used as a framework for the development of certain components in the organization. For example, variable naming standards was derived from this model; Survey Standards used Part 1 for metadata registry, Part 4 for formulation of definitions.

  Geospatial metadata standards used: - (See above)

Accessible and usable geospatially enabled statistics

- International principles or agreed national policies on open data.

  South Africa subscribes to open data policies, within the prescripts of the Promotion of Access to information Act and Copyright Act.

- Accepted standards or guidance on the use of *geographies* for dissemination, visualisation and analysis.

  Various structured tabulating and visualizing tools (like SuperCross, SuperMap) are made available. For any other tabulations the guidance material and metadata is supplied.

- Web service standards.

  The OGC standards Web Map Service and Web Feature Service are used.

- National privacy laws and agreed privacy principals.

  (See above)