



# GIS AND THE 2020 CENSUS

---

Modernizing Official Statistics

AMOR LAARIBI • LINDA PETERS

Cover credits: Income variable data provided by Michael Bauer Research GmbH.

Esri Press, 380 New York Street, Redlands, California 92373-8100

Copyright 2019 Esri

All rights reserved.

23 22 21 20 19 1 2 3 4 5 6 7 8 9 10

Printed in the United States of America

Library of Congress Cataloging-in-Publication Data

Library of Congress Cataloging-in-Publication Data

Names: Laaribi, Amor, author. | Peters, Linda.

Title: GIS and the 2020 census : modernizing official statistics / Amor

Laaribi, Linda Peters.

Description: Redlands, California : Esri Press, 2019.

Identifiers: LCCN 2018060051 (print) | LCCN 2019007152 (ebook) | ISBN

9781589485051 (electronic) | ISBN 9781589485044 (pbk. : alk. paper)

Subjects: LCSH: Census--Methodology. | Census--Geographic information systems.

Classification: LCC HA179 (ebook) | LCC HA179 .L33 2019 (print) | DDC

352.7/50285--dc23

LC record available at [https://urldefense.proofpoint.com/v2/url?u=https-](https://urldefense.proofpoint.com/v2/url?u=https-3A_lccn.loc.gov_2018060051&d=DwIFAg&c=n6-cguzQvX_tUlrZOS_4Og&r=qNU49__SCQN30XC-f38qj8bYYMTIH4VCot2)

[3A\\_lccn.loc.gov\\_2018060051&d=DwIFAg&c=n6-cguzQvX\\_tUlrZOS\\_4Og&r=qNU49\\_\\_SCQN30XC-f38qj8bYYMTIH4VCot2](https://urldefense.proofpoint.com/v2/url?u=https-3A_lccn.loc.gov_2018060051&d=DwIFAg&c=n6-cguzQvX_tUlrZOS_4Og&r=qNU49__SCQN30XC-f38qj8bYYMTIH4VCot2)

[Jb8fvjUA&m=N9LDm816lpUOSrnLbGNLKNgc9sSQOOxCGez3LloTHM&](https://urldefense.proofpoint.com/v2/url?u=https-3A_lccn.loc.gov_2018060051&d=DwIFAg&c=n6-cguzQvX_tUlrZOS_4Og&r=qNU49__SCQN30XC-f38qj8bYYMTIH4VCot2)

[s=IU2vJ0BdojtLU\\_0AhLGaxkrJaFhjH2348dHKCFbJIUA&e=](https://urldefense.proofpoint.com/v2/url?u=https-3A_lccn.loc.gov_2018060051&d=DwIFAg&c=n6-cguzQvX_tUlrZOS_4Og&r=qNU49__SCQN30XC-f38qj8bYYMTIH4VCot2)

The information contained in this document is subject to change without notice.

US Government Restricted/Limited Rights: Any software, documentation, and/or data delivered hereunder is subject to the terms of the License Agreement. The commercial license rights in the License Agreement strictly govern Licensee's use, reproduction, or disclosure of the software, data, and documentation. In no event shall the US Government acquire greater than RESTRICTED/LIMITED RIGHTS. At a minimum, use, duplication, or disclosure by the US Government is subject to restrictions as set forth in FAR §52.227-14 Alternates I, II, and III (DEC 2007); FAR §52.227-19(b) (DEC 2007) and/or FAR §12.211/12.212 (Commercial Technical Data/Computer Software); and DFARS §252.227-7015 (DEC 2011) (Technical Data – Commercial Items) and/or DFARS §227.7202 (Commercial Computer Software and Commercial Computer Software Documentation), as applicable. Contractor/Manufacturer is Esri, 380 New York Street, Redlands, CA 92373-8100, USA.

@esri.com, 3D Analyst, ACORN, Address Code, ADF, AML, ArcAtlas, ArcCAD, ArcCatalog, ArcCOGO, ArcData, ArcDoc, ArcEdit, ArcEditor, ArcEurope, ArcExplorer, ArcExpress, ArcGIS, arcgis.com, ArcGlobe, ArcGrid, ArcIMS, ARC/INFO, ArcInfo, ArcInfo Librarian, ArcLessons, ArcLocation, ArcLogistics, ArcMap, ArcNetwork, ArcNews, ArcObjects, ArcOpen, ArcPad, ArcPlot, ArcPress, ArcPy, ArcReader, ArcScan, ArcScene, ArcSchool, ArcScripts, ArcSDE, ArcSdl, ArcSketch, ArcStorm, ArcSurvey, ArcTIN, ArcToolbox, ArcTools, ArcUSA, ArcUser, ArcView, ArcVoyager, ArcWatch, ArcWeb, ArcWorld, ArcXML, Atlas GIS, AtlasWare, Avenue, BAO, Business Analyst, Business Analyst Online, BusinessMAP, CityEngine, CommunityInfo, Database Integrator, DBI Kit, EDN, Esri, esri.com, Esri—Team GIS, Esri—The GIS Company, Esri—The GIS People, Esri—The GIS Software Leader, FormEdit, GeoCollector, Geographic Design System, Geography Matters, Geography Network, geographynetwork.com, Geoloki, Geotrigger, GIS by Esri, gis.com, GISData Server, GIS Day, gisday.com, GIS for Everyone, JTX, MapIt, Maplex, MapObjects, MapStudio, ModelBuilder, MOLE, MPS—Atlas, PLTS, Rent-a-Tech, SDE, SML, Sourcebook•America, SpatiaLABS, Spatial Database Engine, StreetMap, Tapestry, the ARC/INFO logo, the ArcGIS Explorer logo, the ArcGIS logo, the ArcPad logo, the Esri globe logo, the Esri Press logo, The Geographic Advantage, The Geographic Approach, the GIS Day logo, the MapIt logo, The World's Leading Desktop GIS, Water Writes, and Your Personal Geographic Information System are trademarks, service marks, or registered marks of Esri in the United States, the European Community, or certain other jurisdictions. CityEngine is a registered trademark of Procedural AG and is distributed under license by Esri. Other companies and products or services mentioned herein may be trademarks, service marks, or registered marks of their respective mark owners.

Ask for Esri Press titles at your local bookstore or order by calling 1-800-447-9778. You can also shop online at [www.esri.com/esripress](http://www.esri.com/esripress). Outside the United States, contact your local Esri distributor or shop online at [eurosppanbookstore.com/esri](http://eurosppanbookstore.com/esri).

Esri Press titles are distributed to the trade by the following:

In North America:

Ingram Publisher Services

Toll-free telephone: 800-648-3104

Toll-free fax: 800-838-1149

E-mail: [customerservice@ingrampublisherservices.com](mailto:customerservice@ingrampublisherservices.com)

In the United Kingdom, Europe, the Middle East and Africa, Asia, and Australia:

Eurospan Group

3 Henrietta Street

London WC2E 8LU

United Kingdom

Telephone 44(0) 1767 604972

Fax: 44(0) 1767 6016-40

E-mail: [eurospan@turpin-distribution.com](mailto:eurospan@turpin-distribution.com)

# Contents

Foreword vii

Preface xi

Contributions and acknowledgments xiii

## Chapter 1: Introduction 1

Rationale, scope, and purpose 2

Importance of the integration of statistical and geospatial data 5

## Chapter 2: The role of GIS in census 9

Planning the census process with GIS 10

Needs assessment 16

Success factors 22

What is the GSGF? 25

The GSGF principles—goals and objectives 25

Considerations for rapid changes in technology 30

Esri World Population Estimate 32

## Chapter 3: Changing technology 37

Five key trends in GIS 38

Other key trends 43

## Chapter 4: Establishing the foundation—the geodatabase 51

What is a geodatabase? 52

Spatial data structures 53

Spatial data models 56

Geospatial database design 59

Database modeling 59

Defining census geography 65

Enumeration areas 70

Geographic coding 72

Case study: Albania 76

## Chapter 5: Imagery 85

Introduction 86

Key imagery use cases for national census programs 87

Support for key enumeration business processes 88

Image-processing techniques that support the enumeration process 88

Imagery characteristics important for NSOs 91

Sources of imagery content for national imagery programs 94

Implementation 97

Case study: Portugal 99

## Chapter 6: Planning your work with GIS 103

EA delineation process 104

Maps and sources required 106

Conversion of existing data (analog/digital) 108

Use of imagery in updating EAs 109

Data quality and metadata 119

Creating map services and mobile map packages in support of census operations 127

Use of GIS for optimizing the site placement of field offices 129

Use of GIS for optimizing capital and asset distribution 130

EA optimization 132

Case study: Republic of the Philippines 134

## **Chapter 7: Use of geospatial tools during enumeration 141**

- Fieldwork benefits of using GIS/GPS and imagery 142
- Field data collection 145
- Case study: Jordan 150
- Field operations management 152
- Optimizing workloads 153
- Optimizing routes 154
- Monitoring the progress of census operations 156
- Project management oversight and identifying trouble spots 157
- US Census: Mobilizing field operations 158
- Case study: Arab Republic of Egypt 159

## **Chapter 8: Post-enumeration—processing 165**

- Updating and verifying changes to the database 166
- Aggregation 168
- Database archiving and maintenance 170
- Open platform and system interoperability 171
- Case study: Ireland 173

## **Chapter 9: Post-enumeration—dissemination and analysis 179**

- Dissemination strategies 180
- Level of data release (geographic) 182
- GIS role in creating products for dissemination 194
- Case study: Nepal 205
- Case study: Canada 209

## **Chapter 10: Marketing geographic products and data 217**

- Marketing communication plans 218
- 2020 US Census—future data dissemination and integrated communication plan 219

## **Chapter 11: Contemporary ideas 225**

- ArcGIS® Hub 226
- Smarter maps 228
- Spatial analysis 229
- GeoAI 231
- Geoblockchain 232

## **Appendix: Acronyms 237**

## **Index 241**









A detailed topographic map of the Worb region in Switzerland. The map features contour lines indicating elevation, with peaks reaching up to 947m and 986m. Major roads are shown in grey, including the A6 and A10. Rivers and streams are depicted in blue, such as the Aare and its tributaries. Numerous villages and towns are labeled, including Hub, Ferenberg, Stettlen, Boll, Utzigen, Worb, Enggistein, Rüfenacht, Allmendingen b. Bern, Vielbringen b. Worb, Beitenwil, Trimstein, Richigen, Ried, Nest, Schlosswil, Grosshau, and Walkringen. The map also shows various smaller roads and landmarks like the Klinik Sibah.

## Chapter 1

# Introduction

## Rationale, scope, and purpose

*GIS and the 2020 Census: Modernizing Official Statistics* supports the transformation of countries' censuses with the use of geographic information systems (GIS) and related geospatial technologies to improve data collection, analysis, and dissemination and to enable agencies to build accurate, authoritative, actionable data. The book is specifically intended to provide an up-to-date reference and guide for using geospatial methodologies and techniques to support census and statistical operations.<sup>1</sup> A comprehensive user-oriented guide, the book addresses the needs of census geographers, planners, and managers at all the stages of a census and caters to the needs of the various users of statistical and geospatial products and services.

The increasing use and application of modern geospatial technologies<sup>2</sup> worldwide, particularly during the upcoming round of censuses, have opened a new chapter in all phases of population and housing censuses, improving the efficiency in pre-enumeration, enumeration, and post-enumeration phases. This statement can be attested by Esri through its involvement in many census mapping projects and confirmed by the United Nations (UN) *Principles and Recommendations for Population and Housing Censuses*, the outcomes of a global consultation for the 2020 Round of Census. Rapid changes in technology have brought about improvements in many areas such as computing power, storage, speed, and analytics. Recently, a paradigm shift in the acquisition, resolution, storage, analysis, and distribution of imagery has brought raster analytics to the forefront of projects, offering exponential advances in speed, accuracy, and cost reductions. Digital transformations such as these will have a huge impact on the work of the National Statistical Office (NSO).

A new pattern emerged in conducting censuses during the last decade—a pattern that has ushered in the recognition of the central role of geography to census operations.<sup>3</sup> Geography is increasingly acknowledged as key to virtually all national statistics systems, providing a structure for collecting, processing, aggregating, managing, analyzing, sharing, and disseminating data. In its digital form, geospatial information is exchanged more rapidly, used and reused at will, duplicated without alteration, and easily disseminated to end users. In addition, this geospatial data needs to adhere to the wide variety of security and privacy concerns that are also at the heart of most censuses.

The UN recommended the necessity for countries to keep abreast of technological advances made since the previous round, especially in GIS and Global Positioning Systems (GPS). The UN also recommended that, for the 2020 Round of Censuses, the adoption of GIS should be a major strategic decision.<sup>4</sup> The UN Statistical Commission (UNSC) and the UN Committee of Experts on Global Geospatial Information Management (UN-GGIM) recognized the crucial importance of the integration of statistical and geospatial information for decision-making and have recently recommended the development of a Global Statistical Geospatial Framework (GSGF).<sup>5</sup>



In recognition of this important recommendation, the UNSC and the UN-GGIM began to address the challenges of managing and effectively integrating geospatial and statistical information nationally and globally. They established in 2013 the Expert Group on the Integration of Statistical and Geospatial Information (EG-ISGI). The Expert Group was tasked with developing and advancing the implementation of a GSGF as a basis for the integration of statistical and geospatial information and supporting the building of a statistical-geospatial infrastructure, especially in the context of the 2030 Agenda for Sustainable Development. This importance was also recognized by UN-GGIM regional committees, where it is considered a key priority.<sup>6</sup>

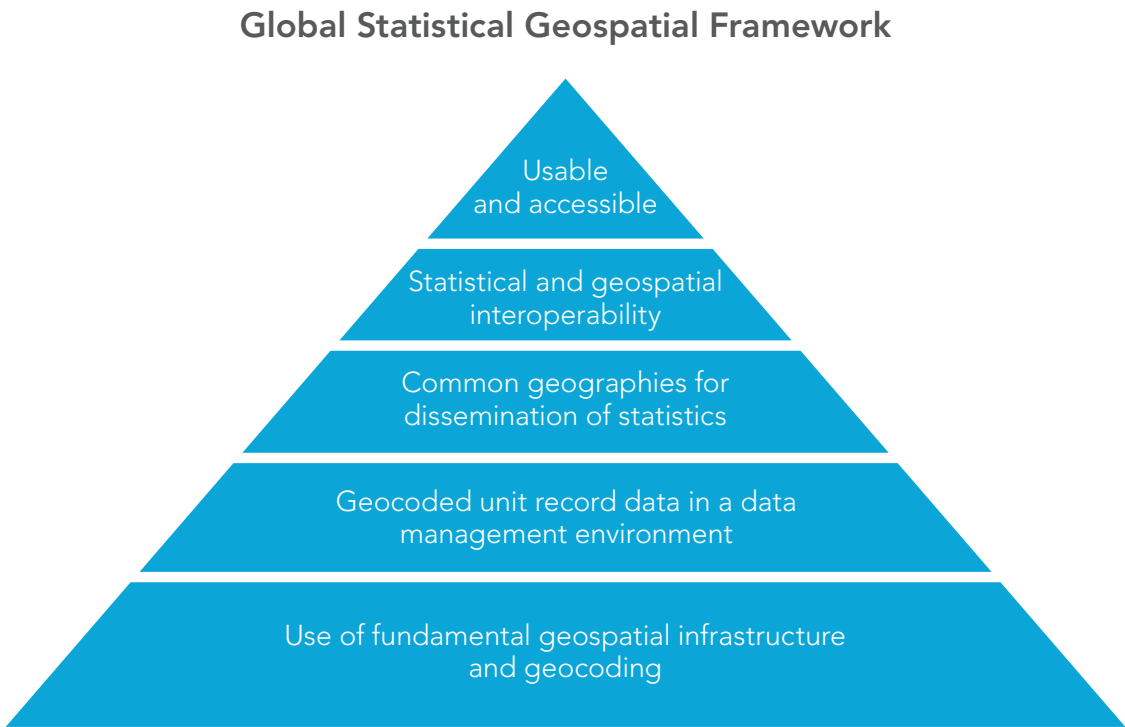


Figure 1.1. The Global Statistical Geospatial Framework (GSGF).

These recent technological developments have already been adopted by many NSOs, including some developing countries. The wealth of national experiences demonstrates that the use and application of GIS is without exception beneficial to the efficiency and quality of the population and housing census. It is understood, however, that a single solution will not fit all national circumstances; hence, there is a need to present various options with guidance to NSOs, allowing them to choose the appropriate solution for their national and local conditions or policies. This book provides guidelines for NSOs to make informed decisions on appropriate GIS technology and how to efficiently use GIS in statistical business processes, taking into consideration different procedures, workflows, organization capabilities, budgets, census timelines, security and application needs, and best practices.

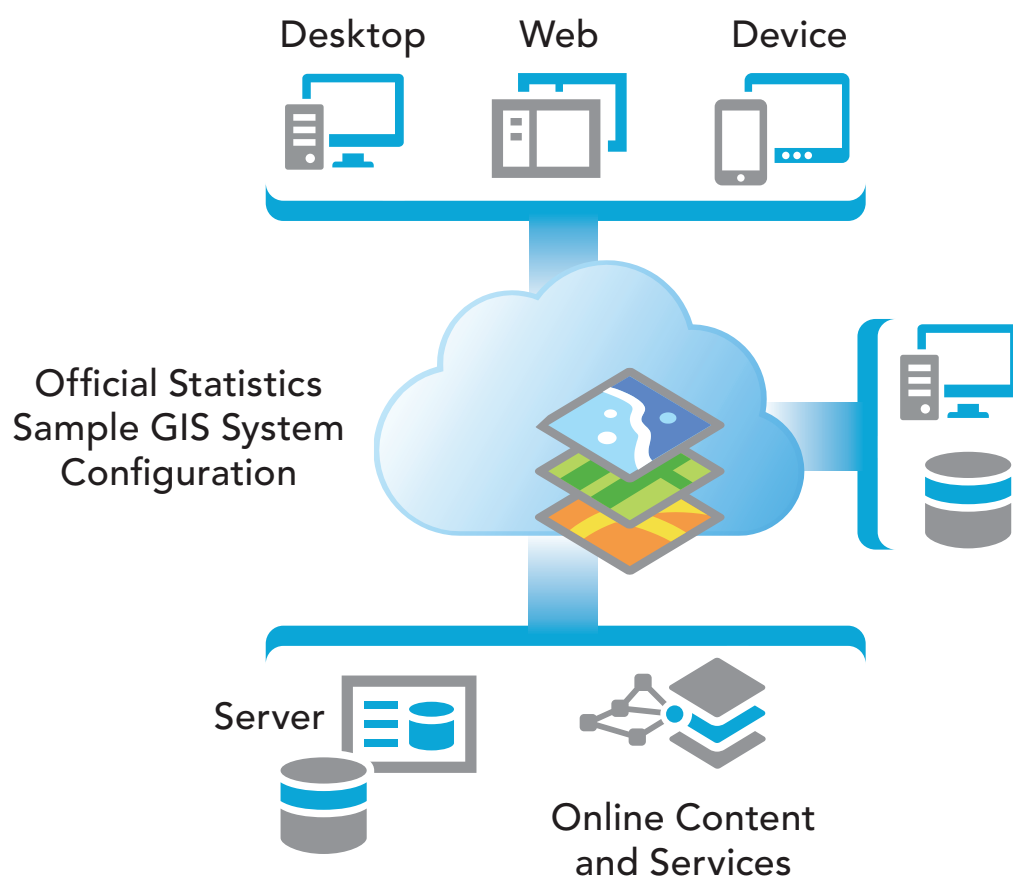


Figure 1.2. Example of a GIS with various types of connections (web, desktop, and mobile).

# Importance of the integration of statistical and geospatial data

The rapid development of modern geospatial technologies has created unprecedented opportunities for geospatial data, and GIS has drastically increased the quantity and quality of geospatial data being collected from a variety of sources. For example, users have access today to imagery that would have been expensive and limited in the past.<sup>7</sup> Geospatial information is becoming a major economic activity with applications across many sectors, bringing increased utility to these sectors and impacting business decisions (Foresman and Luscombe 2017).

With the increasing complexity of national and global challenges, the need to understand interrelationships across the economic, social, and environmental dimensions is crucial for sustainable development at all levels, and the integration of information by geography is critical to designing holistic solutions for any locale. The demand is growing for information on small geographic areas to monitor the development goals and indicators at local and community scales. The integration of statistical and geospatial information has become a significant way to unlock new insights that would otherwise not have been possible by looking at socioeconomic or geospatial data in isolation. It is now understood that the integration of geospatial data with statistics provides significant benefits for nations, as opposed to one-dimensional traditional approaches.

For example, such integration has proved to be critical in achieving improved operational readiness and responsiveness to disasters. By using satellite imagery, scientists and demographers can compare images and statistics taken before and after earthquakes or flooding to estimate the amount of aid to be allocated to populated areas or select safe places in which to temporarily settle the affected population. Using statistical analysis and statistical mapping to study the patterns of urban or rural poverty is a well-established way to provide useful, policy-relevant insights into the patterns of social exclusion and deprivation for local governance policymakers (Baud et al. 2009). Many such examples are related to the increased use of geospatial data in socioeconomic, demographic, and environmental analysis.

Knowing where people and things are and their relationship to one another provides context and improves understanding, aiding in evidence-based decision-making. The geocode is considered one of the basic uniting elements in statistics to which characteristics of persons, households, and enterprises can be attributed.

In this digital-economic age, geospatial technologies are expanding the power of geography and location, from the limited geographical proximity object of Tobler's First Law of Geography to a richer research frontier, enabling us to explore the economic and social impacts of location and spatial relationships. This power is leading geographers to cogitate about a Second Law of Geography for a spatially enabled economy, allowing us to understand why spatial enablement increases functional utility across sectors (Foresman and Luscombe 2017).

*The First Law of Geography states that "everything is related to everything else, but near things are more related than distant things."*

**Waldo Tobler**

It is worth noting that the integration of statistical and geospatial information should be seen in the context of a wider integration of data from a variety of sources. This integration reflects the trends of today's data-driven society through the big data revolution, characterized by satellite imagery, unmanned aerial vehicle (UAV) data, GPS, mobile technology and GIS surveying, volunteered geographic information (VGI) and crowd-sourced data, real-time data streams, cloud-based computing, open data, the Internet of Things (IoT), and geoblockchain. Most studies have stated that spatially enabled data represents at least eighty percent of all data generated within the big data revolution (Foresman and Luscombe 2017).

In today's global community, there is a clear recognition of the need to integrate geospatial and statistical information;<sup>8</sup> the challenge being faced is how best to achieve this integration in an effective and consistent way. It was noted that many NSOs are already transforming, or are planning to transform, their statistical infrastructure, offering an opportunity to embed geography into their national systems and processes. NSOs have been moving from the census mapping approach, limited to the use of GIS for some mapping and dissemination operations, to a GIS-based approach where geography is the foundation of the census, with the use of GIS and other core geospatial technologies at all stages. There is an increasing recognition that the adoption of a geographic-based approach with full integration of statistical and geospatial information has an important role in the modernization of statistics.<sup>9</sup>

GIS can help NSOs realize benefits in terms of reduction of cost and time required, accuracy and quality improvements, and efficiencies in collecting, processing, and communicating information, which all greatly increase the return on investment.<sup>10</sup>

The initial chapters of this book consider the role of GIS in the census, including planning and establishing a full digital GIS-based census program, as well as the importance of geodatabase development. GIS topics relevant to pre-enumeration and enumeration, including the benefits of a digital approach in using GIS throughout statistical business processes, are also addressed. More specifically, for data collection purposes, the book addresses how in a census NSOs need to look at the finest level of geography possible (dwelling and unit level) and capture information at that level. Other chapters focus on post-enumeration GIS tasks and the use of digital mapping in presenting, analyzing, and disseminating census data and marketing geographic products and data. A final chapter presents next-generation trends, focusing on big data and geanalytics.

The book provides statistical organizations with the contemporary methodologies to address user needs at all stages of a census during the 2020 round. Readers will learn how to plan and carry out work using GIS in all phases of a census.

---

## References

- Baud et al. 2009. "Matching Deprivation Mapping to Urban Governance in Three Indian Mega-cities." *Habitat International*.
- GI for Census by the Program Review of the 2010 Round of Censuses carried out by the US Census Bureau for the UNSC in 2015.

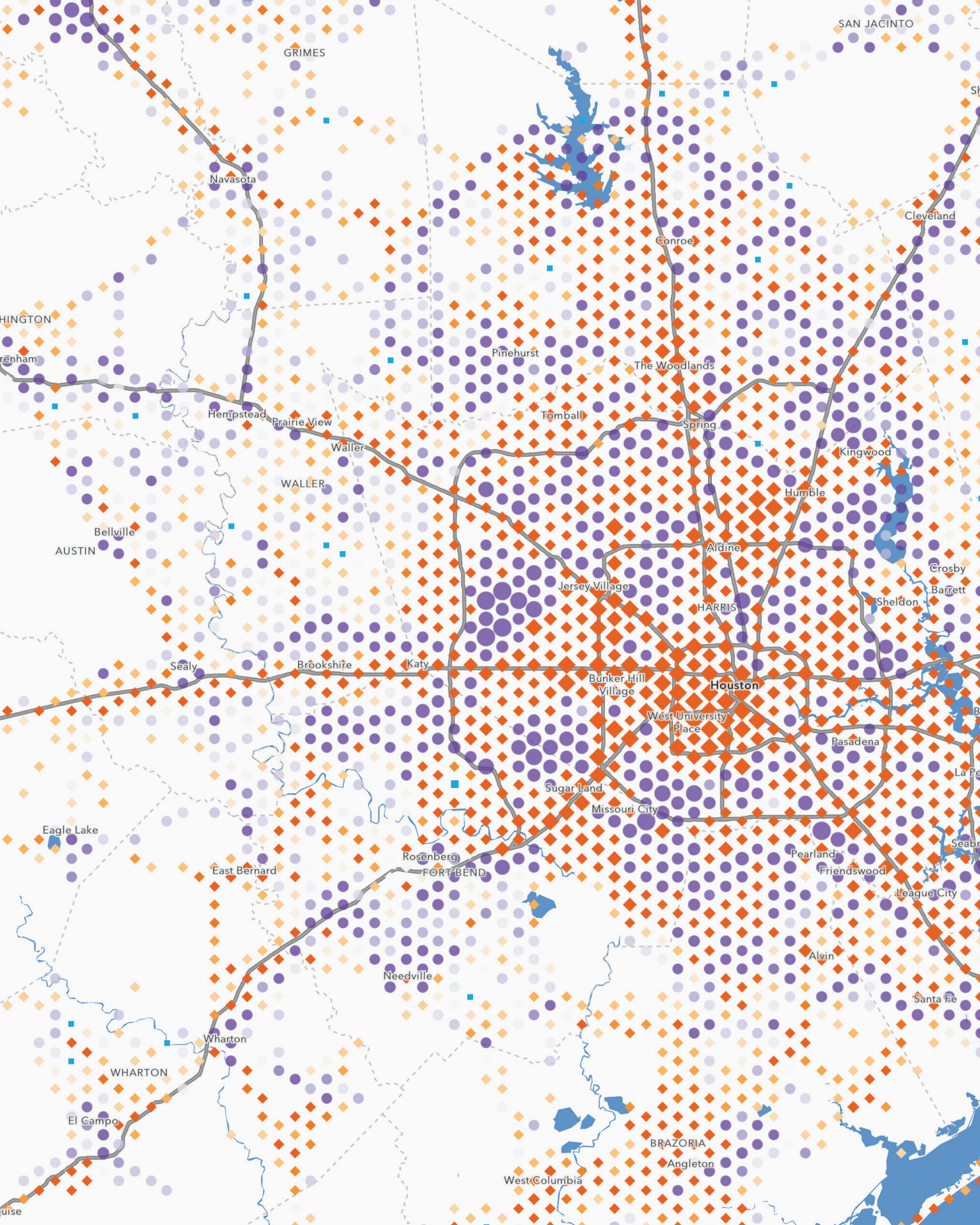


- Report on *National Statistical and Geographical Institute of Brazil: Global Geographic Information Management*, presented at the forty-first session of the UN Statistical Commission. 2010. E/CN.3/2010/13. Available at <https://unstats.un.org/unsd/statcom/doc10/2010-13-Brazil-GGIM-E.pdf>.
- Tim Foresman and Ruth Luscombe. 2017. "The Second Law of Geography for a Spatially Enabled Economy." *International Journal of Digital Earth*.
- UN Economic Commission for Africa/African Centre for Statistics (ACS). 2010. *The 2010 Round of Population and Housing Censuses: A Guide on Census Planning and Enumeration for African Countries*.
- UN Expert Group on the Integration of Statistical and Geospatial Information. 2016. *Background Document on Proposal for a Global Statistical Geospatial Framework*. Available at <http://ggim.un.org/docs/meetings/Global%20Forum/Summary-report%20of%20the%20Global%20Forum.pdf>.
- United Nations Statistics Division. 2017. *Principles and Recommendations for Population and Housing Censuses*, rev. 3. New York: United Nations Publication. Available at [https://unstats.un.org/unsd/publication/seriesM/Series\\_M67Rev3en.pdf](https://unstats.un.org/unsd/publication/seriesM/Series_M67Rev3en.pdf).
- US Census Bureau. 2015. *New Technologies in Census Geographic Listing: Select Topics in International Censuses*. Available at <https://www.census.gov/content/dam/Census/library/working-papers/2015/demo/new-tech-census-geo.pdf>.

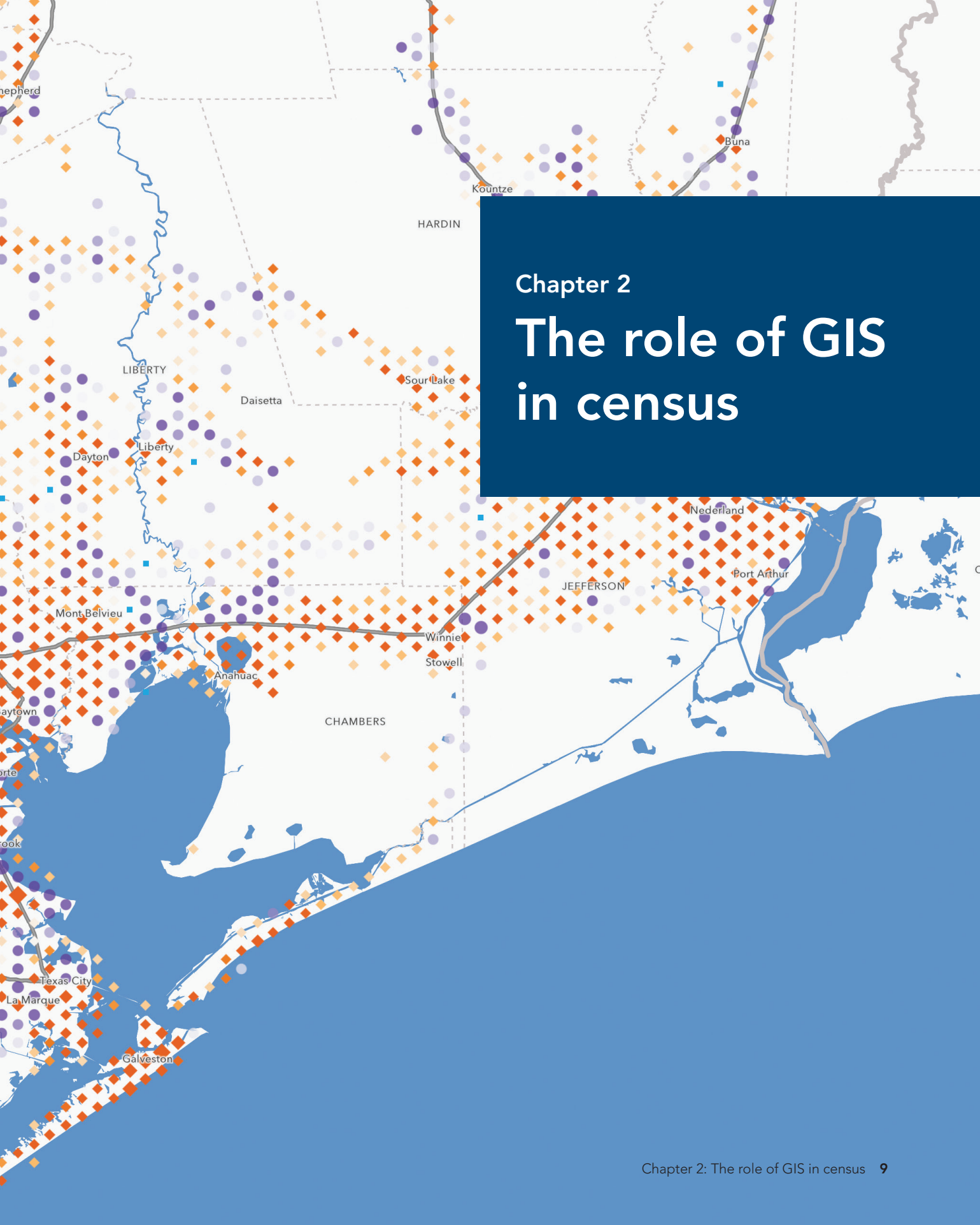
---

## Notes

1. While the main focus of the book is on field-based census approaches and operations, similar GIS techniques can be applied in register-based censuses as well.
2. Geospatial technologies refer to all the means used for the measurement, analysis, and visualization of features or phenomena that occur on earth. They include three core technologies that are all related to mapping features on the surface of the earth: remote sensing (RS), Global Positioning System (GPS), and GIS.
3. The objective of a census is to "count everyone, count them once only, and count them in the right place," according to Preston Jay Waite (US Census Bureau), former associate director for decennial census, US Department of Commerce.
4. See *Principles and Recommendations*.
5. See details on the GSGF in chapter 2.
6. See details at <http://ggim.un.org/regional-entities/>.
7. See *National Statistical and Geographical Institute of Brazil: Global Geographic Information Management*, a report presented at the forty-first session of the UNSC (E/CN.3/2010/13) in February 2010. Available at <https://unstats.un.org/unsd/statcom/doc10/2010-13-Brazil-GGIM-E.pdf>.
8. See *Principles and Recommendations*, para. 3.49: "To ensure complete integration of statistical and geospatial information."
9. See the UN Economic and Social Council's *In-Depth Review of Developing Geospatial Information Services Based on Official Statistics: Note by the United Kingdom Office for National Statistics*. CES. ECE/CES/2016/7.
10. See UNECA/ACS' *The 2010 Round of Population and Housing Censuses: A Guide on Census Planning and Enumeration for African Countries*.







Chapter 2

# The role of GIS in census

# Planning the census process with GIS

## Census geography program

In most countries in the world, a population and housing census is conducted periodically, at least once every ten years in accordance with the UN *Principles and Recommendations*.<sup>1</sup> A population and housing census is considered the greatest democratic operation because it is conducted at the individual or household level, providing information on the main characteristics of a country's entire population in terms of size, geographic distribution, and demographic, social, economic, housing, and living conditions.<sup>2</sup>

Mapping is generally recognized as one of the most crucial activities of a census, playing a critical role in providing the geographic basis used during the actual process of enumeration. Owing to recent technological developments in GIS and other geospatial technologies, the scope of census mapping has been extended to census data analysis and dissemination, with greater efficiency in data collection and enumeration. However, there are some challenges in opting for a full digital census-mapping approach; the use of these technologies has crucial impacts during mapping activities, data collection, processing, analysis, dissemination, evaluation, and archiving, and should therefore be taken into account at an early stage of census planning. Embarking on a GIS-based census should be planned and implemented in a consistent and timely manner.

GIS should be considered an integral part of the census process and carried out continually as a long-term strategy rather than a set of short-term stand-alone mapping and dissemination operations (figure 2.1).

A GIS-based census program aims to achieve the following objectives:

- Support the census planning process.
- Support fieldwork, field operations, and operations management.
- Improve the efficiency and accuracy of the data collected.
- Contribute to analysis.
- Contribute to the dissemination of the census data.
- Integrate statistical and geospatial information for data analysis and evidence-based decision-making and for future censuses and surveys.<sup>3</sup>

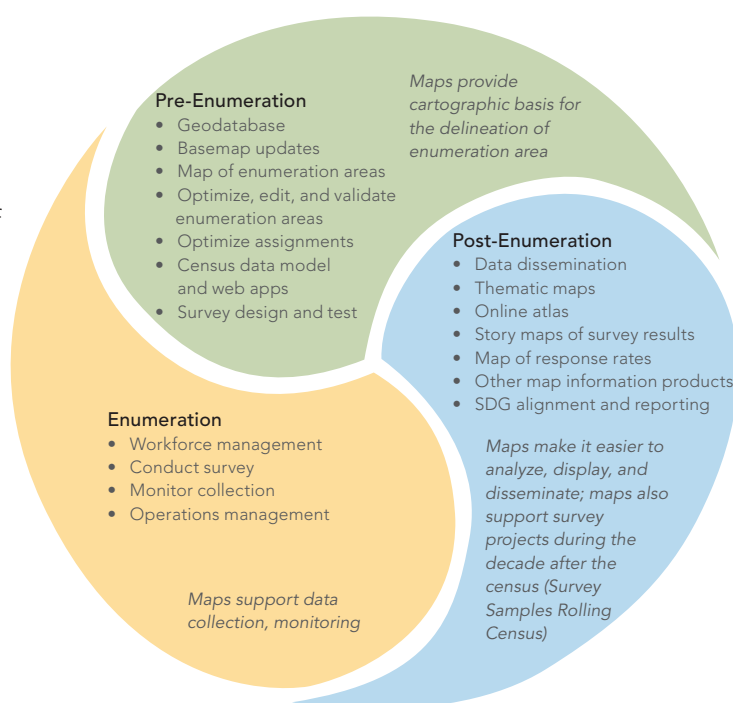


Figure 2.1. The evolving GIS role in census phases.



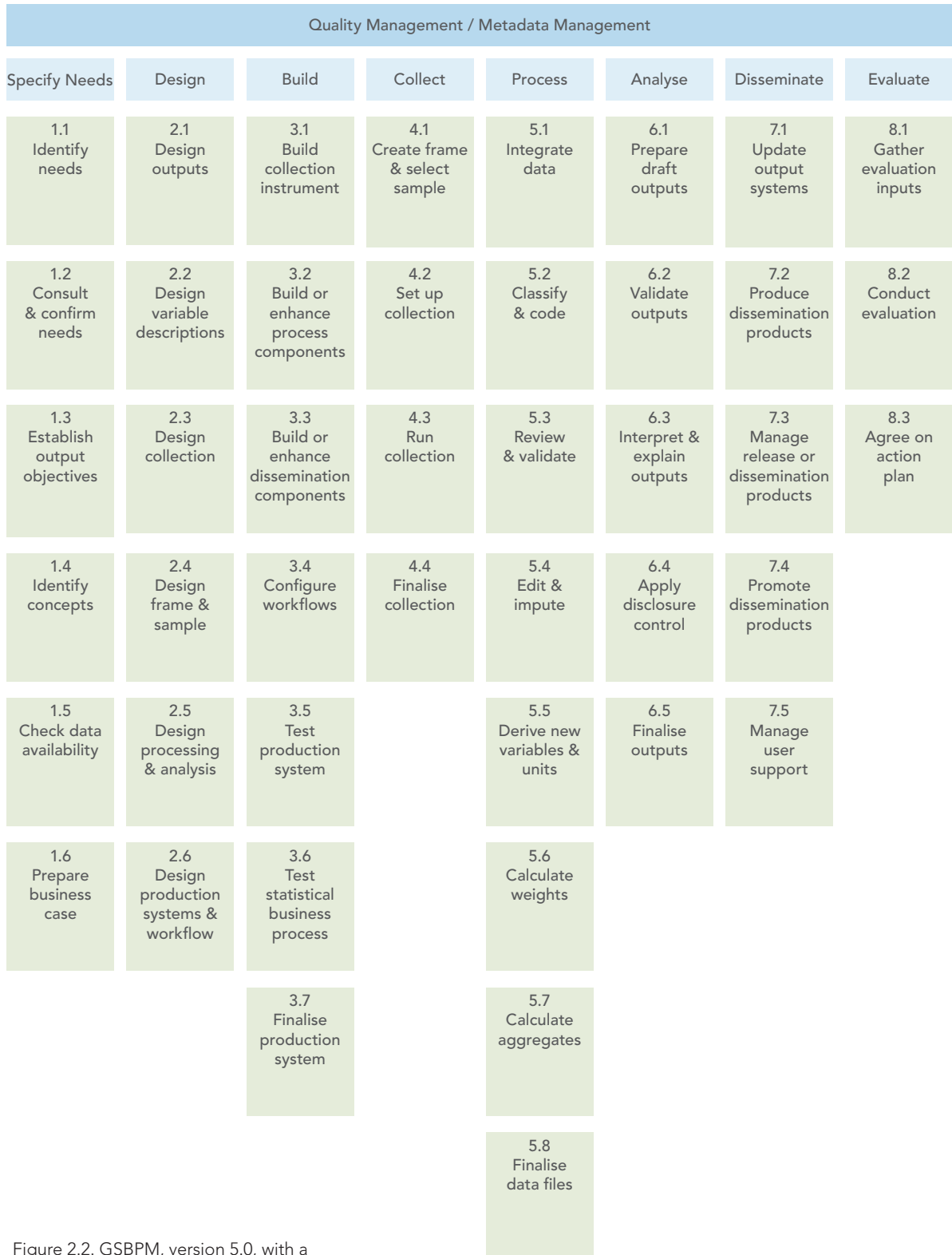


Figure 2.2. GSBPM, version 5.0, with a description of phases and subprocesses.<sup>4</sup>

## Supporting the census planning process

GIS can be applied across the Generic Statistical Business Process Model (GSBPM), shown in figure 2.2. One of the primary objectives of a GIS-based census program is to support the census planning process.

The support includes maximizing coverage while avoiding coverage errors such as noninclusion or double inclusion of units that may result in undercounting or overcounting, respectively. It is important to define explicitly the census geography in terms of hierarchical subdivision of the whole territory into administrative, geographic, and statistical areas, including enumeration areas (EAs) and groups of EAs under the responsibility of supervisors. These geographies are used to estimate the staffing and materials needs and logistics requirements and define the operational zones (or census management areas) for the data collection.

As recommended by *Principles and Recommendations*, the GIS-based census program should be developed at an early stage of census planning to allow sufficient time to produce full national coverage maps (including map services and mobile map packages, which are covered in chapter 6) well before the census date and before the initiation of field-training exercises. Developing a timetable for the GIS and mapping activities is of paramount importance; it's a time-bound operation with the critical date being the date that all enumeration-related maps and map services must be made available to the census field enumeration. This requires the identification of the technical, operational, and institutional tasks to be carried out through the planning process. These tasks include the evaluation of available geographic and technological resources and the critical design issues that determine the nature of the census GIS, with a focus on its core geospatial database and the range of applications

that it will support. Of importance is the inventory of existing data, maps, and other geographic data sources such as imagery. Data conversion and integration processes should also be understood—all of these depend on a well-designed environment and a well-planned operational strategy.<sup>5</sup>

The stages for planning geographic work for the census are shown in detail in figures 2.3a and 2.3b.

The planning stages are divided here into institutional issues, such as the user needs assessment, the determination of the GIS strategy, and the scope of census-mapping activities. The stages also include technical issues such as the explicit definition of census geography, the design of the geospatial database, and the development of clear protocols for data collection. These stages can be carried out more or less simultaneously using organizationally approved methodologies, and many of the choices depend also on the chosen data-integration strategy.<sup>6</sup>

These planning steps may seem difficult to follow or be considered costly by some NSOs. However, once a fundamental geospatial infrastructure is established and GIS is integrated into the statistical business process, the efficiencies gained and benefits realized far outweigh any initial costs. Each country needs to have guiding policies in place to tailor the technological options according to its needs and availability of resources. To justify the investments involved, the strategies of the pre-enumeration mapping program by many NSOs are designed to produce the geographical outputs suitable not only for the census enumeration but also for the operational requirements of data collection activities, dissemination activities, and other user applications such as demarcation of electoral boundaries, defining catchment areas of various public services, monitoring and reporting

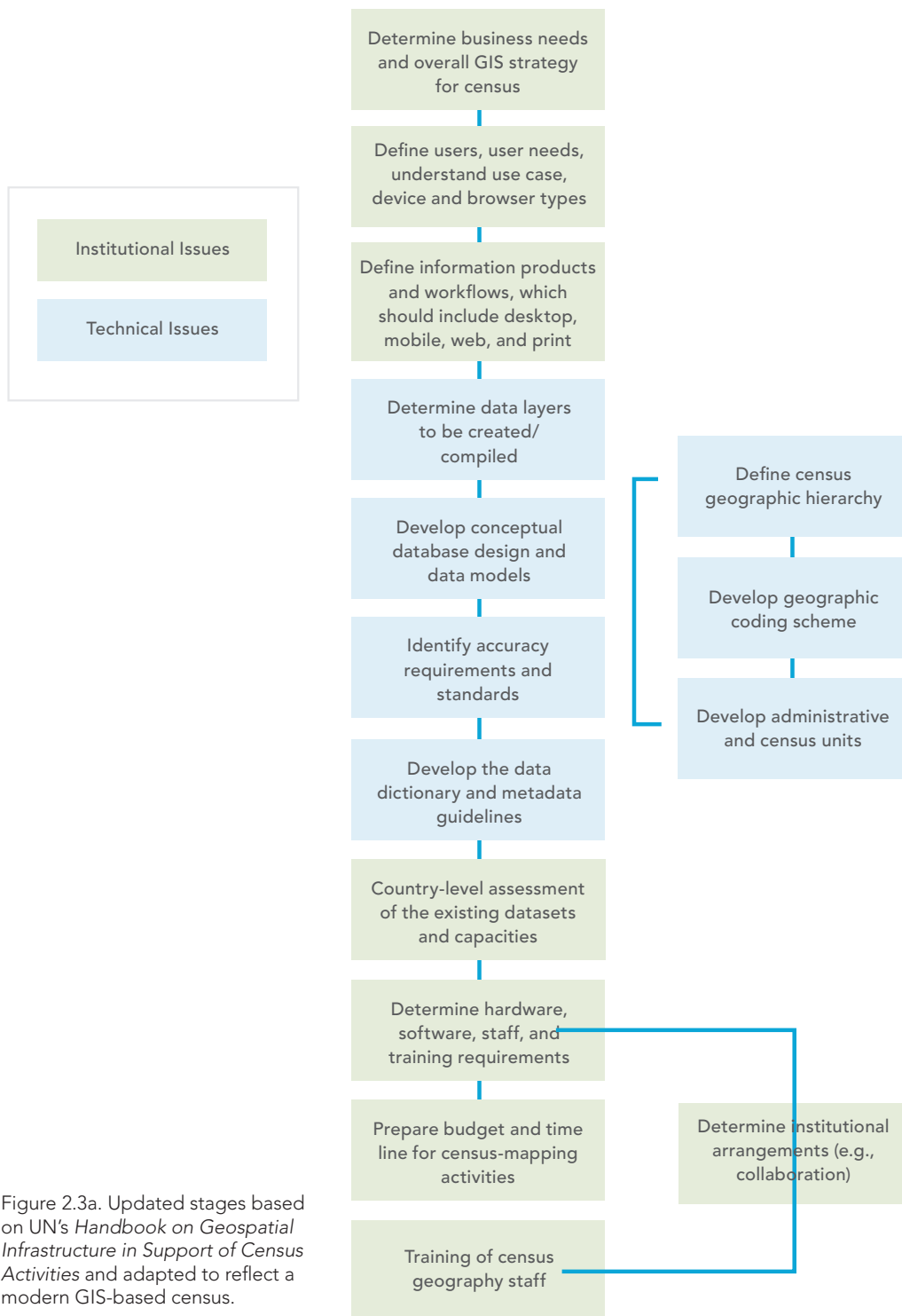


Figure 2.3a. Updated stages based on UN's *Handbook on Geospatial Infrastructure in Support of Census Activities* and adapted to reflect a modern GIS-based census.

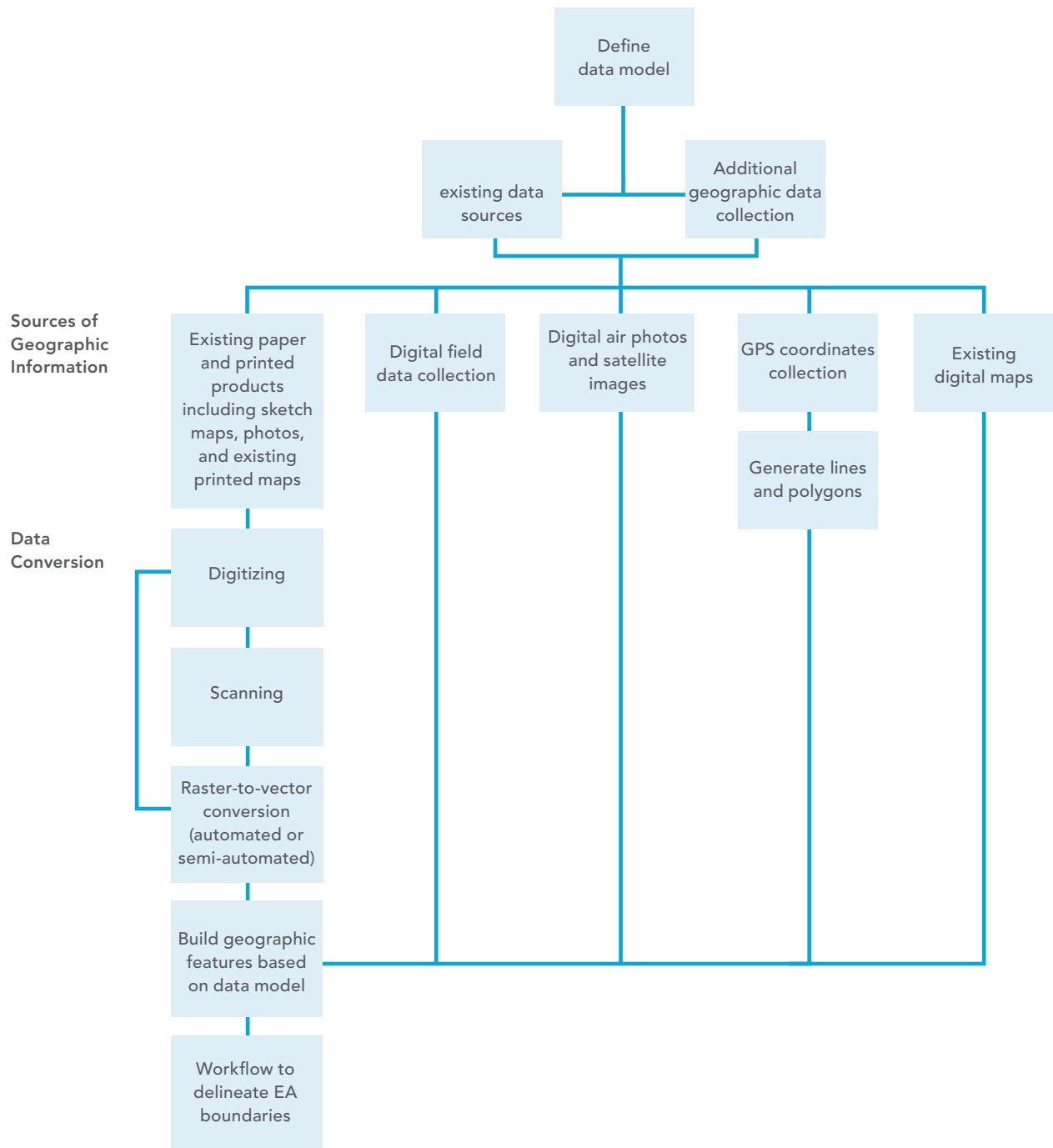


Figure 2.3b. Census planning stages—additional technical issues.



on the Sustainable Development Goals (SDGs), and more. To maximize return on investment in GIS, NSOs should actively participate with other national authorities, including the National Mapping Agencies (NMAs), in establishing a National Spatial Data Infrastructure (NSDI).

### GIS planning considerations

The shift from limited census-mapping operations to an entirely GIS-based digital approach that spans across all stages of the census requires the implementation of an enterprise GIS. The planning and implementation of an enterprise GIS for the census should include a cost-benefit analysis to justify the long-term investment in building a GIS infrastructure. It should also require a comprehensive user-needs assessment; lessons learned from many GIS projects show that a GIS is successful when it provides outputs that meet well-defined user

needs. The user-needs analysis should also assess the available technology in the market and allow the evaluation of different GIS software options to define the most suitable GIS capabilities for census purposes. More specifically, understanding the user needs requires understanding the appropriate hardware, software, and related integrated systems that will provide the specific digital products and services needed. Those products and services can range from interactive ad web-based mapping applications, map services, map packages, cloud computing, and mobile apps to spatial analysis and advanced analytics.

GIS-based planning considerations include the preparation of a detailed implementation plan that corresponds to the different operational phases of the census process. In line with the GSBPM and its equivalent for geospatial applications (see figure 2.4), such a plan should address issues related to

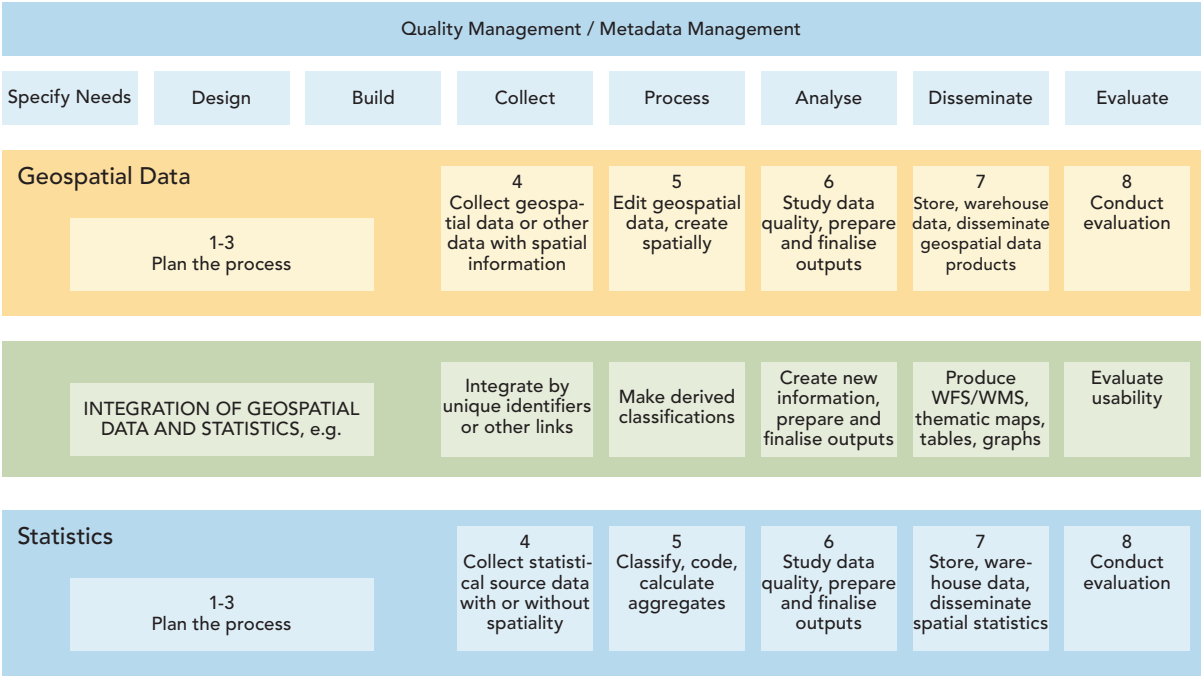


Figure 2.4. The GSBPM and its equivalent dimensions in the production of geospatial data. Source: A Point-based Foundation for Statistics—Final Report from the GEOSTAT 2 Project.

the design of the geospatial database, data conversion and integration, and all the other phases of geospatial data production.

Because GIS was not traditionally considered part of the core business of the NSO, awareness workshops about the benefits of GIS should be organized to win executive support for the program. Technical trainings and workshops about the use of GIS should be provided to the staff involved in the census and supporting functions. In addition, the plan should also involve workshops for the major data users about the potential of GIS and draw on institutional mechanisms to mobilize cooperation among all stakeholders to establish a GIS infrastructure at a national level.

GIS can be leveraged to support a wide variety of activities across the entire census process. Following are some examples of how GIS is being applied today in census preparation. These activities will create the foundation for all census work to follow:

- Creation or updates to geodatabase and basemaps.
- Creation or updates and validation of EAs.
  - Production of digital EA maps for fieldwork and operations.
  - Use of remote-sensed data in EA creation or update.
  - In-office address canvassing can be done using imagery where applicable and where quality data is available.
  - Where in-office validation is not possible, maps can be used in field verification or *in situ* validation.
  - Integrating fieldwork using remote-sensed data.
- Conducting GIS analysis to ensure complete and balanced coverage. Overlaying EA maps on a

scaled national basemap to ensure the absence of omissions or duplications.

- Applying GIS analysis to facilitate efficient census operations.
  - Using GIS analysis to determine the most efficient placement of field offices.
  - GIS-based analysis to optimize allocation of field-workers to EAs based on various criteria such as language, distance, hours of work, and more.
- Creation of map services (the way maps are made available to the web).
- Creation of mobile map packages (the way maps are made available on a mobile device in an off-line mode).

More details are provided in the following sections and subsequent chapters.

## Needs assessment

It is generally recognized that a well-defined plan that involves various users with different data needs should be prepared at an early stage. Using geospatial information technology in all the stages of a census requires even more focus and the need to identify, understand, and plan how to address the specific GIS and mapping needs and expectations of the main user groups, as well as the census geographic products required.

The needs assessment plan should include strategies on the content to be discussed during the consultations with users, questions on the ways census products will be accessed (device type) and used as well as their presentation, and the kind of training that needs to be provided to users.

Conducting a comprehensive needs assessment with a business need-driven approach rather than a technology-driven approach helps to define

user needs. The assessment should also identify any available resources within the NSO (and in the country), such as maps suitable for the census operations, existing software packages and related equipment, qualified staff in GIS, and financial resources for the GIS-based census program. This identification will assist in reconciling user expectations with what is feasible given available resources, working backward from final products and services to requirements.<sup>7</sup>

The next section will elaborate on the needs assessment of the main user groups: (a) major users of census data, (b) persons and institutions participating in the census operations, and (c) the general public and civil society.

### **Major users of census data**

A user-oriented census provides major data users—governmental departments, local administrations, the academic and research sector, and the private sector—with easy access and clear understanding of the statistics available, enabling them to benefit from census results. One of the major impacts of a GIS-based census is the extension of the community of users, interested in statistics with a geographic dimension. The assessment of needs in this case includes consultation on demographic and geographic content desired, geographic structures such as administrative hierarchies or geographic units needed for data collection or data aggregation,

and geographic base products (maps, imagery, and other remotely sensed data) that support analysis and dissemination of census data. A consultation of the needs of these major users is necessary to determine the form and scope of data to be disseminated and to ultimately understand their expectations in terms of census geographic data products.

### **Persons and institutions participating in the census operations**

Evaluating the needs for census-mapping activities is equally important for the persons (employees and contractors) and institutions participating in the census operations. To obtain an understanding of existing resources and requirement gaps, the NSO must carry out a survey of available human resources, hardware, and software, and must conduct an inventory of existing data (in digital format and on paper) and any requirements for data conversion. The NSO should also understand any ongoing or planned relevant GIS-oriented activities by other public and private entities. The purpose of this assessment is to avoid duplicating efforts, which is key to reducing the cost of census geographic operations and delivering census products on time.<sup>8</sup>

Conducting a comprehensive needs assessment to answer some of the critical questions shown in the following table will help the NSO consider which new technologies to adopt to modernize the census process.

Figure 2.5. Critical questions that NSOs should consider for a needs assessment.

Critical questions	Example areas to be considered
<b>What strategy will be adopted for the use of geo-spatial technology?</b>	Will the agency use GIS technology throughout the statistical process and across the organization? Or will it be used only in certain functional areas (e.g., if GIS is adopted it can serve as the base for future surveys beyond population and housing such as agriculture, labor force, etc.), thereby improving return on investment for the organization?
<b>What are the costs involved?</b>	What are the costs required to adopt the geospatial technology, including costs for new hardware, upgrades of hardware, or existing software and any data needed? This should include an understanding and consideration of ongoing needs to maintain the systems whether purchased outright or built in-house.
<b>What advantages are expected to be realized?</b>	Identify the benefits that the use of geospatial technology will bring to the census (e.g., improved accuracy, efficiencies in operations, cost savings, or delivery of data in a timelier manner).
<b>What are the costs or benefits of implementing this new system or method for the agency?</b>	Proceed with a cost-benefit analysis to justify the GIS-based method.
<b>What is the ease or difficulty of installing and maintaining hardware or software?</b>	Ease of installation and maintenance are concerns for many NSOs because they may not have technical expertise on staff to support. NSOs may need to consider outsourcing or cloud-based deployments, for example.
<b>Are both the software and the database scalable? To what extent can the system manage peak loads?</b>	The performance of both the GIS software and the geospatial database should be scalable: able to grow and be used for multiple purposes. Appropriate system sizing and scaling will need to be done to consider peak upload periods across all systems.
<b>What type of users will access the system?</b>	Some types may include employees, field supervisors, area managers, executives, contractors, citizens, and application developers. How will their identity be defined, and what will they have access to?



Figure 2.5. Critical questions. (continued)

Critical questions	Example areas to be considered
<b>What types of accessibility?</b>	<ul style="list-style-type: none"> <li>• What type of device(s) will be used to access?</li> <li>• What functionality is required or allowed by the user?</li> <li>• Will workers be both online and off-line?</li> </ul>
<b>What kind of mapping, GIS, and imagery capabilities are needed?</b>	<ul style="list-style-type: none"> <li>• Do executives, field supervisors, or area supervisors (for example) need different tools?</li> <li>• Are there plans to use a street basemap or an image basemap or both?</li> <li>• What scale or resolution is required?</li> <li>• What geographic reference layers (infrastructure, natural environment) are needed?</li> </ul>
<b>What other data should be considered?</b>	<ul style="list-style-type: none"> <li>• How will administrative data be used?</li> <li>• Are geographic names or searches required?</li> <li>• What point of interest (POI) data is needed (hospitals, schools, etc.)?</li> </ul>
<b>What measures should be taken for geocoding?</b>	<ul style="list-style-type: none"> <li>• What level of accuracy is needed in geocoding?</li> <li>• What geocodes will need to be captured, updated, and used?</li> </ul>
<b>What support will be required?</b>	<ul style="list-style-type: none"> <li>• Documentation and reference manuals.</li> <li>• Help and vendor support: <ul style="list-style-type: none"> <li>◦ Availability of local support from the vendor to provide any needed support in a reasonable amount of time.</li> </ul> </li> </ul>
<b>What are the staff training needs?</b>	<ul style="list-style-type: none"> <li>• Is this existing workforce adequately trained? If not, can training be provided in time frames needed?</li> <li>• More specifically: <ul style="list-style-type: none"> <li>◦ Are staff trained in the production of maps and cartographic databases?</li> <li>◦ Are staff trained in the creation of map services?</li> <li>◦ Are staff trained in basic GIS concepts, implementation, and data management?</li> </ul> </li> </ul>

Figure 2.5. Critical questions. (continued)

Critical questions	Example areas to be considered
<b>What are additional key considerations about EA maps, devices, imagery, and GPS?</b>	<ul style="list-style-type: none"> <li>• Device needs should include GPS and built-in maps based on current available digital geography.</li> <li>• Device or application needs to allow for edits to census geography in the field.</li> <li>• Device needs to allow for secure capture of survey data in the field in both online and offline modes.</li> <li>• EA maps should be designed for ease of use by enumerator.</li> <li>• EA maps should be designed for efficiency.</li> <li>• Use of satellite imagery to verify boundaries and identify new buildings and housing locations.</li> <li>• Use of GPS to georeference points including housing location and important landmarks.</li> </ul>
<b>What security needs should be considered?</b>	Can the GIS integrate with existing security systems and does it allow for industry-standard security principles and controls to be applied at all levels?

### Cost-benefit analysis

An important component of needs assessment is to proceed with a cost-benefit analysis. A consideration of the costs and benefits of using GIS (and other geospatial technologies) across all stages of the census is essential for making informed decisions about how much to invest and showing the benefits from GIS use. The costs, as stated earlier, include the acquisition of equipment, software, maintenance, services, hiring additional staff or training existing staff, acquiring data from outside the NSO, and data conversion costs. Benefits can be achieved in many areas including time and cost savings, improvements in staff effectiveness and efficiency, improvements in accuracy, and the earlier dissemination of finished products.

Country experiences have shown that the benefits derived from the use of GIS include productivity gain and time savings; cost savings; cost avoidance;

greater credibility and authority of geographic and map products and survey data; better service; increased accuracy; improved consistency; income generation; improved analysis; improved policy-making; improved data sharing; and improved outreach.<sup>9</sup> While measuring all the expected benefits can be difficult because they are only partially quantitative, their impact in terms of improvement of data accuracy and data quality—the ultimate purpose of a census—is recognized by countries that use a comprehensive, large-scale GIS in their censuses.

### General public and civil society needs

Rapid changes in technology, combined with the growing demand by the public and civil society for data and information access and transparency, are shaping the needs of users in the community. In some cases, it is necessary to provide specific

data or access to meet the needs of these users, as individuals or communities. Community engagement is key for organizations as producers of official statistics. Citizens and civil society alike need to fundamentally understand the issues of the day and where change is needed, which comes from NSO authoritative data.

Modern geospatial platforms provide easy-to-use tools (e.g., apps) for data sharing and collaboration, which allow for even more citizen engagement. NSOs may need to educate the general public about the power of geospatial tools in conveying census information to users, and how spatial analysis related to GIS capabilities allows users to analyze and visualize data, leading to a more meaningful interpretation of results. For example, census data is essential and most relevant in public safety efforts. NSOs must also keep abreast of and anticipate the general public demand for their data, including the need for data to support the understanding and reporting on the UN SDGs (see the case study on Ireland and SDGs in chapter 8).

### **Information product needs**

The ultimate purpose of conducting a comprehensive needs assessment of the three user groups is to determine the range of output or information products that need to be completed both throughout and at the end of the GIS-based census program. This determination allows the NSO to rethink the geographic products from the users' perspective, and how to make them available in a form suited to their needs supplemented by proper and useful documentation, including metadata.

Generally, the geographic information products include digital EA maps for the creation of aggregate dissemination areas; geographic boundary files in digital format that provide statistical reporting units used for census dissemination and spatial

analysis; a road network file in digital form representing the country's national road network and containing information on street names and types for all major urban areas;<sup>10</sup> and correspondence files that indicate how current reporting units relate to those used in previous censuses and facilitate a comparison between sequential censuses (see the online ancillaries at [esri.com/Census2020](http://esri.com/Census2020)).

Other information products may include the list of changes to municipal boundaries showing the

## **Volunteered geographic information**

Public demand for information is expanding rapidly, as is "citizen science"—the scientific research conducted, in whole or in part, by amateur scientists. A widespread community engagement called volunteered geographic information (VGI) is popularizing the creation, sharing, and use of geospatial data, often through smart phones, social media, and online mapping tools, and acting as a valuable mechanism to encourage public participation and citizen engagement in geospatial information. VGI, crowdsourced data, and social media may provide essential information to citizens and local knowledge to communities on their neighborhoods, health facilities, zoning conditions, or even the damage from flooding or some other local disaster. Citizens and particularly some focus groups may need statistical data to combine with field data to create community mapping. Even though the quality or accuracy of such mapping may not be high, the speed of collection and mapping of the data can be performed easily, making it useful for an early response.

changes brought to municipal boundaries, status, and names that occurred between the current and previous census; a set of vector layers containing feature data, such as landmarks, roads, schools, health facilities, etc., which can be used for population-based spatial analysis; centroids/point-based files (point data based on GPS coordinates or individual addresses) that provide respectively geographic point reference for each reporting unit and other geographic features (such as dwellings, places, etc.); and gazetteers that provide geographic coordinates and naming for all population settlements and other important geographic entities in the country (such as administrative units, places, and census units).<sup>11</sup> Modern systems also allow for the creation of operational dashboards which can be useful during the full census cycle, giving transparency and increasing communication and collaboration across teams. Finally, in dissemination, many different types of information products may be required, including maps, story maps, atlases, online web applications, and more (more details are provided in chapter 9).

## Success factors

Besides conducting a comprehensive needs assessment that helps adequately define the user needs and identify the available resources within the NSO and in the country, particularly the funding requirements, we need to consider critical factors to succeed in a full digital GIS-based census program. Chief among these factors are: (1) ensuring commitment from senior management for building a long-term digital program; (2) building technical and human capacities required for sustaining the GIS-based systems and databases and setting up an independent unit for cartography/GIS activities within the NSO; (3) using technical standards; (4)

developing a partnership for cooperation with the NMA and other organizations involved in geospatial information activities; and (5) choosing the appropriate methodology of integration of the new geospatial technologies with the census-mapping operations (compatibility). These factors are findings revealed by survey-based studies and lessons learned from country experiences during the last 2010 Round of Censuses.<sup>12</sup> We will elaborate on them in the following sections.

### Human factors

While a GIS-based census project may be perceived as technical in nature, what ultimately drives its success often hinges on human factors, particularly in terms of ensuring commitment from senior management for building a long-term program and the true integration of geospatial and statistical data. It is critical, as stated earlier, to consider the GIS-based census program as a long-term program that benefits not only the national census office but also other government agencies and ministries as well as census data users. Approaching GIS as an integral part of the process and knowing that it requires an important initial investment can help achieve a return on investment in the long term.

With any new methodology with little critical mass of experiences known by policy makers and good practices confined to specialists, a promoter and a champion are needed. These people will not only steer the GIS-based census process but will also reach out to the decision makers, advocating for the benefits of using this new approach. A champion can communicate and demonstrate the power of GIS to visualize, explore, and analyze data related to development, ensuring appropriate and timely government support for all aspects of the program.



### **IT and human capacity building**

A GIS-based census program carried out as a continuous process requires an independent geographic information unit to be set up within the NSO, with a permanent staff that can carry out the geospatial analysis and census-mapping tasks spanned over all census stages and between censuses. A permanent unit that can develop partnerships, understand and adhere to data standards and interoperability requirements, forge agreements and contracts for data collection and sharing, and cooperate with the other major actors in building the NSDI of the country is a valuable asset for success. In the absence of in-house GIS capabilities, it becomes difficult for the census office to make qualitative checks on the data created and delivered by the contracted companies. This difficulty may result in the delivery of inaccurate and deficient data, which in turn affects the quality of the geospatial data and map products used for the census.

Building technical and human capacities required for sustaining the GIS-based census program is a recurrent issue to be addressed. Building this capacity requires training personnel with appropriate skills in GIS to serve all census geography needs. A well-trained staff is a key factor for the success of GIS-based census-mapping projects. Country experiences have shown that retaining skilled GIS staff within NSOs is a major issue, particularly with the proliferation of numerous GIS applications unrelated to the census and often commercially more attractive and lucrative. To retain trained qualified staff, NSOs need to provide significant incentives and ensure continual training to keep employees abreast of advances in technology.

### **Use of standards**

Standards are important to every industry and GIS is no exception. The benefits of developing, adopting, and implementing technical standards and common metadata have been recognized by both statistical and geospatial communities. Those benefits enable interoperability and facilitate the integration and use of diverse sources of statistical and geospatial data and services in all sectors of a global economy. It is recognized that the development of common standards and metadata would specifically make statistical and geospatial information more usable and relevant to a wider range of users.

Implementing geospatial databases for the census and related geographic products requires the creation and maintenance of metadata guidelines and a data dictionary. The metadata standard and its use beyond the census would increase the usability and relevance of this data and enhance the ability to compare data both within and across geographies. NSOs should adhere to applicable standards established for the country because they are necessary to ease the effort of sharing census data and understanding between organizations and between applications. Many applications extend beyond national borders, including those related to disaster management, spread of infectious diseases, humanitarian assistance, migration, and cross-border displaced population, for which it is necessary to integrate datasets across national borders.

Geospatial standards and metadata are not always available in countries. When they are unavailable, it is recommended that NSOs follow published basic standards and definitions, especially about the geospatial features used in statistical operations, such as boundaries of statistical units, buildings, private dwellings, and collective living quarters.

As stated previously, work led by the UN has resulted in the development of a Global Statistical Geospatial Framework (GSGF) used as a guideline for integrating statistical and geospatial information. An overview of this framework is presented in the box.<sup>13</sup> The development of an overarching statistical-geospatial framework, as part of an overall information architecture at both the national and global levels, is needed for enhancing the common understanding of concepts such as geocoding, georeferencing, and other terminology; sharing common standards; and fostering collaboration between both statistical and geospatial communities. This recognizes that when statistical and geospatial information is integrated within a statistical-geospatial framework, the geostatistical results can significantly improve the quality of official statistics and population and housing censuses and support the measuring and monitoring of the SDGs. This also explains and justifies the shift from census mapping to GIS-based approaches, with full use of GIS and geospatial technologies across all stages of the census and beyond.

## What is the GSGF?

The GSGF, developed under the auspices of the EG-ISGI, is a high-level framework that consists of five broad principles that are considered essential for integrating geospatial and statistical information (see figure 2.6).

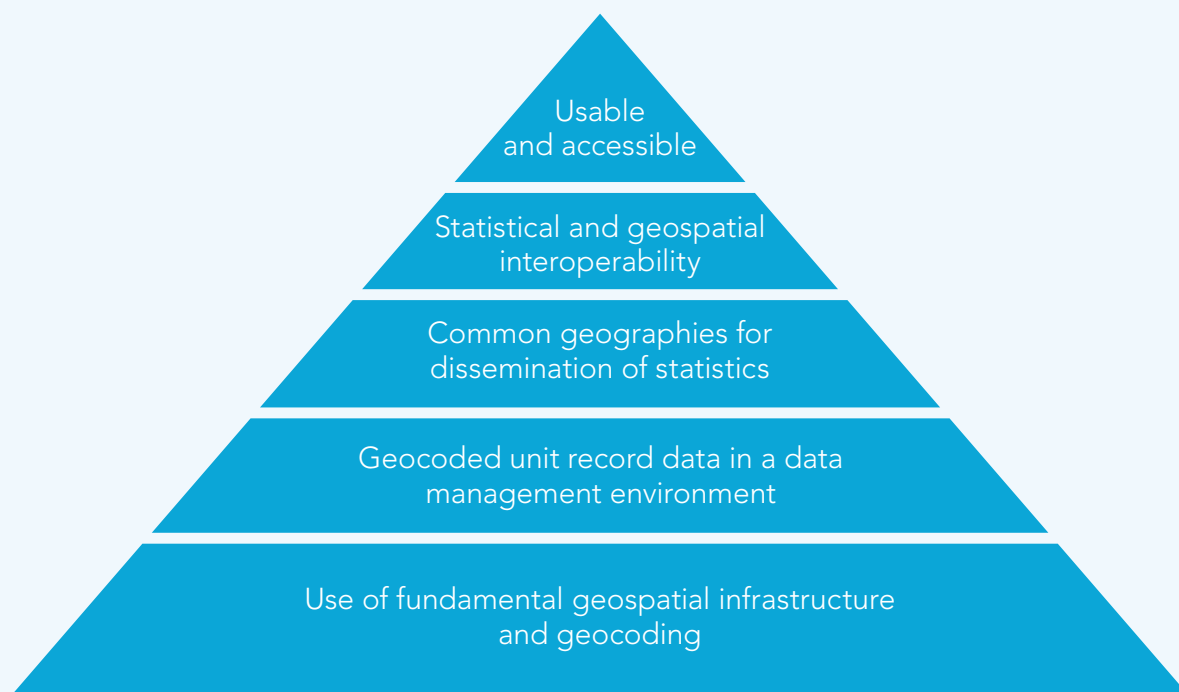


Figure 2.6. GSGF—five fundamental principles.

## The GSGF principles—goals and objectives

Each of the high-level principles\* in the GSGF is defined by a set of goals and objectives and is supported by international, regional, and applicable national standards and best practices. At its sixth session in 2016, UN-GGIM adopted the following five principles:<sup>14</sup>

### Principle 1: Use of fundamental geospatial infrastructure and geocoding

This foundational principle is about the fundamental geospatial infrastructure. In this context, it refers to those geospatial datasets that underpin the core business of NSOs. The requirement is about “a common and consistent approach to establishing the location and a geocode for each unit in a dataset, such as a

---

\* This text explaining the principles is mainly derived, with some adaptations, from the background paper *Proposal for a Global Statistical Geospatial Framework*, available at <http://ggim.un.org/meetings/GGIM-committee/documents/GGIM6/Background-Paper-Proposal-for-a-global-statistical-geospatial-framework.pdf>. Also used is the *Global Statistical Geospatial Framework: Linking Statistics and Place* document, available at [ggim.un.org/meetings/GGIM-committee/8th-Session/documents/Global-Statistical-Geospatial-Framework-July-2018.pdf](http://ggim.un.org/meetings/GGIM-committee/8th-Session/documents/Global-Statistical-Geospatial-Framework-July-2018.pdf).

person, household, business, building or parcel/unit of land. . . . The goal of this principle is to obtain a high quality, standardized physical address, property, 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).” The fundamental geospatial infrastructure will differ between countries, but “the process of obtaining locations and geocodes should use relevant, fundamental geospatial data . . . from the National Spatial Data Infrastructures or other nationally agreed sources.” In some nations, it is likely to include national address lists, land parcel databases, road networks, business registers, and building footprints. In other nations, it is likely to include national and regional boundaries, town and locality locations, etc. However, the EG-ISGI recognizes that geocoding of statistical units using point referencing is highly preferable when compared with merely associating statistical units with a geographic region (i.e., a polygon).

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

This principle focuses on the requirement that the linkage for each statistical unit record should occur within a data management environment to allow all statistical data to be applied to any geographic context (i.e., in aggregating data into a variety of larger geographic units), including future changes to geographies over time, and enable data-linking processes that aim to integrate information of varying nature and sources by so-called linked data techniques. This principle is underpinned by two combined factors: (1) key to geospatial enablement and flexibility, the finer the better; and (2) statistical data management ensures confidentiality. While principle 1 is about geocoding infrastructure, principle 2 is about applying this geocoding infrastructure to unit records and ensuring that the statistical infrastructure can use geospatial infrastructure and standards for geocoding.

### **Principle 3: Common geographies for dissemination of statistics**

This principle is about the enablement of comparisons across datasets from different sources, which requires using a common set of geographies for the display, reporting, and analysis of social, economic, and environmental information. The goal of this principle is to provide the tools for integration of aggregate data on common geography and for the dissemination of data. In addition to the recognized importance of traditional statistical and administrative geographies, NSOs should consider the benefits of grid-based statistical systems because gridded data can be both a rich source of information and a consistent geography for disseminating and comparing information.

### **Principle 4: Interoperable data and metadata standards**

This principle aims at greater efficiency and simplification of the creation, discovery, integration, and use of geospatially enabled statistics and geospatial data; increasing the potential application of a larger range of data and technologies; and a wider range of data available and accessible for use in comparisons and analysis in decision-making: “Both the statistical and the geospatial data communities operate their own general data models and metadata capabilities; however, these models are often not universally applied.



The statistical community uses the Generic Statistical Information Model (GSIM), the Statistical Data and Metadata eXchange (SDMX), and the Data Documentation Initiative (DDI) mechanisms. The geospatial community uses the General Feature Model (GFM) and the ISO19115 metadata standard, plus a number of application-specific standards.” Within the statistical community, there is a need to build geospatial processes and standards into statistical business processes in a more consistent manner. In consequence, the EG-ISGI has recognised that a top-down approach is required, incorporating geospatial frameworks, standards, and processes more explicitly into the Common Statistical Production Architecture and its components. Work has started internationally on how to better enhance interoperability between statistical and geospatial metadata standards.

### **Principle 5: Accessible and usable geospatially enabled statistics**

This principle is about the ultimate purpose of developing this GSGF: Informing decision-making in a more flexible, agile approach using new and existing data and reassuring custodians of safe release (i.e., release of their data with confidence, privacy, and confidentiality protected) while enabling users with discovery and access of geospatially enabled statistics. More specifically, this will consist of identifying or, where required, developing policies, standards, and guidelines that support the release, access, analysis, and visualization of geospatially enabled information. In addition, data users can undertake analysis and evaluation. Web services enable machine-to-machine access and dynamic linkages of information. As stated in principle 4, this will allow data to be accessed and used efficiently and ultimately support the modernization agenda of many NSOs and NMAs.

The UNSC adopted the framework at its forty-eighth session in March 2017, and subsequently the Committee of Experts endorsed it at its seventh session in August 2017.

The standards and best practices that will form the detailed guidance for countries implementing the framework are still under consideration by the EG-ISGI and will be brought to the UNSC and UN-GGIM for consideration when finalized.

Where standards, policies, or datasets required to support the framework do not currently exist, the framework provides a clear mandate for their establishment. Collaboration between countries and within the 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. A number of areas of further work have been identified and are under consideration by EG-ISGI.

The framework is driven by strong principles but remains broad enough that countries can adapt it to their local conditions and needs. Many members agreed to prepare country-level examples detailing their application of the GSGF principles to their national conditions in preparation for broader global community input. In this regard, a template for country examples for each of the five principles has been prepared by the Australian Bureau of Statistics and available for use by countries.

Other countries such as Sweden are now beginning to share their interpretations on how they will implement the GSGF within their organization.

## Collaboration

As stated previously, NSOs need to be aware that the challenges involved in introducing GIS and other geospatial technologies are not only technical, but more often institutional and organizational. Mobilizing cooperation among all stakeholders and building partnerships right from the beginning of the census-planning phase is essential to implement a GIS-based census program catering to the needs of all major users.

For example, since producing basemaps is not a core competency of the NSO, developing a partnership for cooperation with the NMA can be mutually beneficial: the NMA provides the NSO with basemaps updated and appropriate for the census operations, and, in return, the mapping agency may benefit from the improvements provid-

ed by the fieldwork undertaken by the NSO and the geographic products derived from the census.

Other examples include the selection of GIS data or systems that should be made considering the use by other government departments, allowing that the construction of a national census geospatial database can be employed in many different national contexts for numerous purposes. Institutional arrangements facilitate partnership for data standards and interoperability, agreements and contracts for data collection and sharing, and collaboration across government at a variety of levels including national, regional, and local, thus contributing to the building and development of the national geospatial information infrastructure of the country.



Figure 2.7. Collaboration across agencies and the community.

# Geography and Statistical Data Are Foundational

## An Integrated Data Model Is Essential

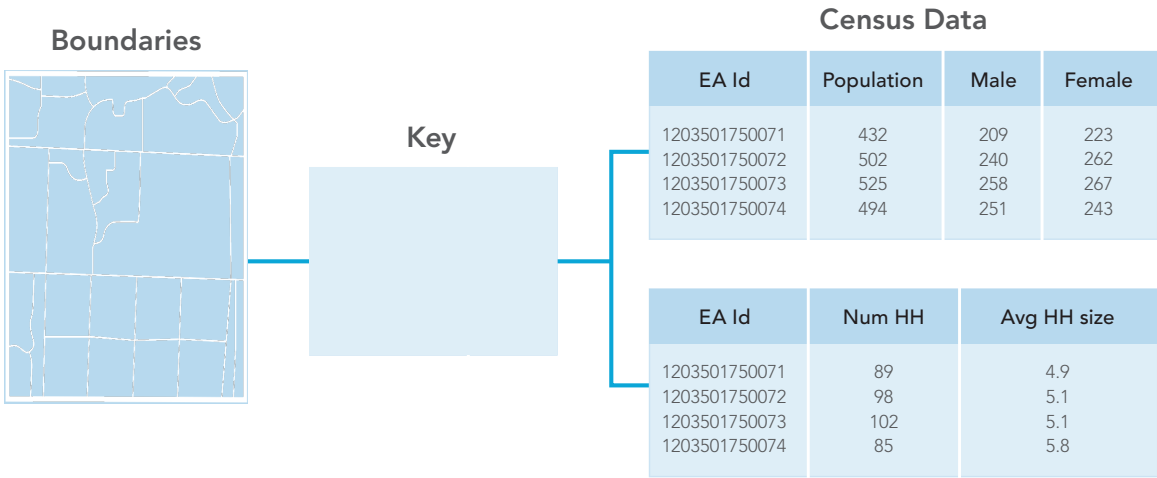


Figure 2.8. Integration of geospatial and statistical data. Statistical data can be easily mapped by using geographic information such as geocodes, EA codes, and EA boundaries.

### Compatibility—methodology of integration

Choosing the appropriate methodology of integration of the new geospatial technologies with the census operations is another central factor in the adequate integration and optimal use of geospatial technologies. Given the different approaches and national experiences, it is generally recognized that no universal solution fits all; in deciding what approach to adopt, the NSOs should be guided by their needs, policies, and availability of resources allocated to the census in their countries.

Lessons learned from different experiences point to challenges in integrating existing census or geo-

graphic data with newly collected data and converting the data into digital form. Consideration should be given to the construction and implementation of a geospatial database, required data models, and maintaining standards, metadata guidelines, and data dictionaries.

For example, the NSO may need to extract, transform, and load (ETL) some data received from other agencies or sources owing to variations in file formats. By considering standards, system requirements, data gaps, and other key issues, the NSO will be better prepared for success.

## Considerations for rapid changes in technology

How do NSOs keep abreast of technology changes and understand what and when to implement? Artificial intelligence (AI), machine learning (ML), IoT, and smart devices hold the promise for an intelligent future. The challenge will be understanding and leveraging these new technologies in both a timely and cost-efficient manner and mapping the NSO's business needs and workflows to these potential new technologies.

Understanding new ideas such as this, in addition to continuous adaptive security or microservices, will be critical to future success. Intelligent applications (apps) and analytics are another key trend we must be prepared for. Intelligent apps have the potential to transform the nature of work and the structure of the workplace. While NSOs are conducting their work today, they must at the same time plan for the future.

The GIS industry is no stranger to this transformation, according to Jack Dangermond, president and founder of Esri: "The geospatial industry is quickly evolving because the capabilities of digital connectedness and collaboration are moving ahead exponentially. In fact, the five biggest trends in the area of geographic information system (GIS) technology are centered on making data more accessible and creating context to visualize this data in an age when fast, easy access to information is taken for granted."



Figure 2.9. Jack Dangermond, president of Esri, addresses a conference audience.





Figure 2.10. Top five trends in GIS technology today.

According to Dangermond, the top five trends in GIS technology today are as follows:

1. Location as a service
2. Advanced analytics
3. Big data analytics
4. Real-time GIS
5. Mobility

Dangermond continues: "The last leap in computing was the shift from the server to the cloud. Software as a service (SaaS) opened a world of opportunities for GIS, as shared map services like the World Imagery basemap are no longer separate from the unique services offered to users. GIS users can share data, collaborate, make mashup maps in the server, and then connect to the cloud.

The next leap in GIS technology and computing is connecting to the vast network of devices providing data in real time. This technology is a revolutionary change and brings great opportunity. The more accessible data is, the more important it will be to understand it. And maps are the visual language for understanding the context of data."

An example of these technologies coming together today is the Esri World Population Estimation model. Chief Cartographer Charlie Frye leads this project on behalf of Esri.

## Esri World Population Estimate



Figure 2.11. Image of the Sinai Peninsula. Courtesy of MDA NaturalVue.

MDA NaturalVue is imagery processed with natural colors (RGB channels) to reflect the earth's true colors. This cloud-free imagery is available even for areas with persistent cloud coverage.

As methods like these become more repeatable and proven, it will influence the way we conduct our own census and survey work.

This project brings together big data, advanced analytics, and location as a service. In December 2014, Esri published the initial version of the World Population Estimate (WPE) image service on ArcGIS Online. The service represented a global estimate of the 2013 population of the earth as a footprint at 250-meter resolution. The WPE is a dasymetric raster surface in which each cell represents an estimated count of people living in the location the cell represents. The same method is applied globally, which means the values

in each of the cells are intended to be comparable. Data used includes Landsat8 panchromatic imagery, classified land cover from MDA Information System's BaseVue® 2013, and a global dataset of road intersections. A texture-detecting model is used to locate settlement in the rural areas beyond where the urban classes from BaseVue 2013 exist. Because land cover data inherently includes information about surface water, agriculture, and permanent ice, some of these factors can be eliminated because there is a high level of certainty that people as a rule do not live in these locations.

Esri's purpose in creating the WPE is to facilitate population-based spatial analysis—specifically, providing a footprint for ArcGIS tools to perform the following:

- Provide a basis for estimating redistributions of other population characteristics with higher geographic fidelity than traditional choropleth mapping techniques. These characteristics include demographic variation, economic behaviors, political attitudes, and cultural characteristics.
- Support cross-boundary estimation of populations that:
  - Are affected by natural disasters and complex humanitarian emergencies (for mitigation, response, and recovery).
  - Live within high-risk environments, such as areas prone to flooding or forest fires.
  - Are exposed to disease (to prevent transmission).
  - Are affected by disease outbreak (to direct resources to support effective treatment efforts).
- Estimate human impact on the environment for purposes of sustainability, resilience, and planning along the lines of green infrastructure.

In summer 2016, Esri published an update representing the estimated footprint and population for 2015. The 2015 estimate included additional image services for population density and confidence. The 2015 services were all based on 150-meter resolution.<sup>15</sup>

In early 2018, Esri published a population 2016 estimate that included an improved confidence score. This work continues, and Esri is now collaborating to further this research in a consortium called POPGRID (<https://www.popgrid.org>). Hosted by the Center for International Earth Science Information Network (CIESIN) at Columbia University, this group works on a voluntary basis to share and document techniques for producing global gridded dasymetric redistributions of population, to foster greater awareness of gridded population estimates, and to produce comparisons of the global gridded products. The POPGRID participants include Oakridge National Laboratory, Esri, CIESIN, the Joint Research Centre (JRC), the European Commission, WorldPop (University of Southampton), the German Aerospace Center (DLR), Image-Cat, Inc., the US Census Bureau, the World Bank, Facebook®, and Google®.

Using imagery and other data sources and analyzing them using machine learning and AI, each of these organizations is working to innovate and create a global mosaic of high-resolution settlement layers that will provide an increasingly higher quality estimate of human population distribution.

---

## References

- Andrea Gros and Tobias Tiecke. 2016. *Connecting the World with Better Maps*. Available at <https://code.facebook.com/posts/1676452492623525/connecting-the-world-with-better-maps/>.
- Charlie Frye. 2015. *Map Gives New Insights into Global Population*. Available at <https://medium.com/esri-insider/map-gives-new-insights-into-global-population-d7b1902f7c15>.
- Charlie Frye. 2016. *Updated Population Dataset Sharpens Focus on the Human Footprint*. Available at <https://blogs.esri.com/esri/esri-insider/2016/03/30/updated-population-dataset-sharpens-focus-on-the-human-footprint/#more-6823>.
- Jack Dangermond. 2017. *Five GIS Trends Changing the World According to Jack Dangermond, President of Esri*. Available at <http://geoawesomeness.com/five-gis-trends-changing-world-according-jack-dangermond-president-esri/>.

---

## Notes

1. Since its inception, the UN recommends that each country conduct a population and housing census at least once every ten years, usually within the so-called census round. The 2020 census round covers the ten-year period 2015–2024.
2. The UN *Principles and Recommendations* defines the population census as “the total process of collecting, compiling, evaluating, analyzing and publishing or otherwise disseminating demographic, economic and social data at the smallest geographical level appropriate pertaining, at a specified time, to all persons in a country or in a well-delimited part of a country.” In the case of a housing census, this same process focuses on “statistical data relating to the number and condition of housing units and facilities as available to the households pertaining, at a specified time, to all living quarters . . . and occupants thereof in a country or in a well-delimited part of a country.”
3. These objectives sum up the findings of the UN regional workshops held worldwide during the 2010 Round of Censuses and presented at the 2015 International GIS Workshop in China.
4. See <https://statswiki.unece.org/display/GSBPM/GSBPM+v5.0>.
5. See details in the UN Department of Economic and Social Affairs Statistics Division’s *Handbook on Geospatial Infrastructure in Support of Census Activities*, 2009. Available at [https://unstats.un.org/unsd/demographic/standmeth/handbooks/series\\_f103en.pdf](https://unstats.un.org/unsd/demographic/standmeth/handbooks/series_f103en.pdf).
6. See *Handbook on Geospatial Infrastructure*.
7. See *Handbook on Geospatial Infrastructure*.
8. See *Handbook on Geospatial Infrastructure*.
9. See *Handbook on Geospatial Infrastructure*.
10. For example, Statistics Canada has its boundary files and the road network files available in three formats: ArcGIS® shapefile (.shp), Geography Markup Language (.gml), and MapInfo® (.tab).
11. See *Handbook on Geospatial Infrastructure*.

12. See details in *Australia Programme Review* and the related background paper on *Developing a Statistical Geospatial Framework in National Statistical Systems: Survey of Linking Geospatial Information to Statistics—Analysis of Questionnaire Responses* (2013), available at <https://unstats.un.org/unsd/statcom/44th-session/documents>, and the UNSD report on the *Results of a Survey on Census Methods Used by Countries in the 2010 Census Round*, available at <https://unstats.un.org/unsd/censusb20/KnowledgebaseArticle10696.aspx>.
13. For details, see the report of the UN Expert Group on the Integration of Statistical and Geospatial Information, 2015, E/CN.3/2015/37, at <https://unstats.un.org/unsd/statcom/46th-session/documents>.
14. See *UN Expert Group on the Integration of Statistical and Geospatial Information Background Document on Proposal for a Global Statistical Geospatial Framework*. Available at <http://ggim.un.org/meetings/GGIM-committee/documents/GGIM6/Background-Paper-Proposal-for-a-global-statistical-geospatial-framework.pdf>.
15. See the Esri story map “We Are Everywhere” available at <https://landscapeteam.maps.arcgis.com/apps/MapJournal/index.html?appid=23d43b031a4d47629c511909188cfefcd>.



# GIS AND THE 2020 CENSUS

Modernizing Official Statistics

**Census workers need to capture and analyze information at the finest geographic level with mobile and geospatial-based technology.**

*GIS and the 2020 Census: Modernizing Official Statistics* provides statistical organizations with the most recent GIS methodologies and technological tools to support census workers' needs at all the stages of a census. Learn how to plan and carry out census work with GIS using new technologies for field data collection and operations management. After planning and collecting data, apply innovative solutions for performing statistical analysis, data integration, and dissemination. Additional topics cover cloud computing, big data, location as a service (LaaS), and emerging data sources.

While *GIS and the 2020 Census* focuses on using GIS and other geospatial technology in support of census planning and operations, it also offers guidelines for building a statistical-geospatial information infrastructure in support of the 2020 Round of Censuses, evidence-based decision-making, and sustainable development. Case studies illustrate concepts in practice.

---

**Amor Laaribi** has a PhD in geomatics sciences and over 30 years of experience in geospatial information for decision-making and development. He joined the United Nations Statistics Division (UNSD) in 2000 to define and plan the priorities of the division in geospatial information and to promote its use with census mapping in developing countries.

**Linda Peters** has a bachelor's degree in geography and over 20 years of experience in the geospatial industry. In her 14 years at Esri, she has consulted with companies across many industries, helping them achieve greater efficiencies through spatial and location analysis. She works with National Statistical Offices (NSOs) across the globe, helping them understand how to apply geographic methods and analysis to census and statistical activities.