New Technologies in Census Geographic Listing

Select Topics in International Censuses[1]

Released December 2015

INTRODUCTION

This Select Topics in International Censuses brief will provide national statistical offices (NSOs) with focused information about technologies for census geographic listing operations that have matured in the past ten years. Geographic listing strongly affects subsequent census operations. During the listing operation, census workers identify dwelling units and list households within the census area, which is usually the entire country. Traditionally, when census workers performed this count using pencil and paper, they created pictographic (sketch) maps for use during the full population and housing enumeration. Leading into the 2000 census round, Geographic Information Systems (GIS) software became widely available on desktop computers with user-friendly graphical interfaces. However, it was during the 2010 census round (2005–2014) that many NSOs transitioned from paper maps to digital. The transition from paper to digital maps involves the digitization of enumeration area boundaries as represented on sketch maps. These sketch maps must also portray physical boundaries for proper digitization. Physical features in digital format that correspond to those represented in the sketch maps are then used to rebuild enumeration area boundaries. During the 2010 census round, many NSOs also began researching and implementing solutions using enterprise GIS, satellite imagery analysis, and Global Positioning System (GPS)-enabled handheld devices. However, cost and complexity formed significant barriers to widespread adoption. For the 2020 census round, NSOs will increasingly adopt these technologies for census operations.

Geographic Definitions Used in This Brief

Each NSO may have its own understanding of the terms used when describing listing operations. These definitions are intended to provide clarity for the purposes of this brief, not to supplant existing NSO definitions or suggest the need for complete uniformity. There are further distinctions between dwellings based on assets and services available. For more information on these definitions, refer to pages 192–197 of *Principles and Recommendations for Population and Housing Censuses* (UNSD, 2015).

Building points—A complete listing of the point locations of structures within the enumeration area—usually the complete area of a country—that may or may not contain one or more living quarters.

Living quarter—Any inhabited structure that is not wholly used for nonresidential purposes at the time of the census.

Dwelling unit—A room or group of rooms (place of abode) within a permanent structure constructed specifically for occupation by one household with access to a street or common space.

Continued on next page

[1] This technical note is one in a series of "Select Topics in International Censuses" exploring matters of interest to the international statistical community. The U.S. Census Bureau helps countries improve their national statistical systems by engaging in capacity building to enhance statistical competencies in sustainable ways.

United States Census Bureau
U.S. Department of Commerce
Economics and Statistics Administration
U.S. CENSUS BUREAU
census.gov
USAID
FROM THE AMERICAN PEOPLE

Continued from previous page

Unconventional dwelling unit—A room or group of rooms (place of abode) that due to their construction: 1) were not expected to be durable, 2) were purposefully intended to be mobile, 3) were constructed for a purpose other than human habitation, or 4) lack services to the extent that they are considered unfit for habitation but are used for that purpose.

Housing units—Comprised of all conventional and unconventional dwelling units; that is, all separate and independent places of abode occupied at the time of the census.

Household—A group of people who participate in some degree of shared decision-making and engage in common planning and preparation for the provision of food.

Enumeration area—The operational unit for data collection during the census that comprises the lowest level of the geographic hierarchy of administrative and statistical units.

Figure 1.
Photo Interpretation Example

A number of objects in this scene could be buildings, but many are not. Photo interpretation relies on visual identification and best judgment.

Source: U.S. Geological Survey.

SATELLITE IMAGERY ANALYSIS

Sketch maps sometimes include building points, but digitization of building points from paper sources rarely provides an acceptable level of locational accuracy. For an accurate digitization of building points, NSO staff can use satellite imagery to capture building points on a desktop computer while handheld GPS technology can be used for field verification. The combination of these two technologies can reduce the resources required for building listing. These techniques make it possible to capture building points rapidly and to work around the lack of an address system. However, distinguishing between living quarters and nonresidential buildings, as well as registering individual housing units, still requires fieldwork.

NSO staff can use imagery acquired by satellites or aircraft to collect building points using photo interpretation techniques and widely available GIS desktop software. Two approaches can be used: 1) every possible structure can be captured or 2) staff can visually interpret the imagery to eliminate objects that are not habitable. For example, a highly reflective object with sharp, linear edges could appear to be a building. However, if an object does not cast a shadow, it is likely dry, packed soil. This type of interpretation is open to error and dependent on the experience and training of the digitization staff (see Figure 1). For that reason, all possible objects can be captured through digitization pending field verification for inclusion in the final dataset.

Imagery analysis can also be automated using tools available in desktop GIS software. Imagery classification, for example, analyzes the characteristics of satellite imagery to create classes based on shared spectral or spatial characteristics. This method is useful for categorizing building points and identifying areas newly settled since the last census. However, any automated classification approach requires multi-band imagery. Each band represents part of the electromagnetic spectrum, allowing for classification analysis. Base maps streamed from an Internet service—such as Google, Microsoft, or Esri—cannot be classified because the images lack spectral data. If an NSO relies on streamed imagery, manual classification is the only available option. Furthermore, while imagery classification can aid in the production of some datasets, such as road networks and newly settled areas, imagery classification does not replace fieldwork for verification of living quarters.

Imagery classification techniques can be divided into two broad categories: per-pixel and per-object. Per-pixel classification refers to the assignment of each pixel—the smallest data unit present in an image—to a category based on how the pixel reflects light. Per-object classification refers to the extraction of objects—components of the landscape (e.g., buildings, roads, fields, fences)—based on their shape and how they reflect light. Object-based classification is not usually found in standard GIS software and must be purchased or developed separately. NSO staff would also require specialized training to use object-based classification software.

Figure 2.
Imagery Classification Techniques and Required Staff Skills

	Land Cover Techniques accessible to advanced GIS users. Change in software may increase training requirements.	**Land Activity/Land Use** Project lead must have specialized training and remote sensing research experience.
Pixel-Based Possible using standard GIS software.	Change detection of new settlements and urban clusters.	Change detection of areas within and adjacent to existing settlements; estimate housing units.
Object-Based Requires specialized software.	Change detection of new buildings and physical features (roads, fences, power lines).	Possible to produce per-building maps with estimate of number of living quarters and housing units.

Source: U.S. Census Bureau.

A further distinction exists between classification of land cover and classification of land use. Imagery classification to determine land cover is a well-established form of geospatial analysis and is available in most GIS and imagery analysis software. Advanced GIS users can usually be trained to perform useful land cover analysis. However, land use classification is more advanced, largely experimental, and requires additional geospatial layers—such as zoning and cadastral maps—that may not yet exist in a country. Conducting land use studies requires specialized training experience not typically found at NSOs (Figure 2).

The current state of image classification technology can complement the creation and maintenance of geospatial data. However, such software is unlikely to replace the need for time-intensive manual digitization when using imagery for address listing purposes in the near future.

FIELD GIS DATA COLLECTION USING HANDHELD DEVICES

Handheld devices have been in use for geographic listing activities since the 1990s[2] when field portable GPS technology became widely available. These devices have contributed to the improved accuracy of digital GIS data.

Current purpose-built GPS devices can collect attributes accompanying geographic data, but may not support the design complexity needed during a census geographic listing operation without extensive modifcation. Handheld GPS devices are also limited to features that are accessible. Linear features such as rivers and roads present challenges since they must be easily navigable by field staff. Collecting boundaries with handheld GPS devices presents additional difficulties when clear physical markers on the ground are absent.

The emergence of tablets and smartphones that incorporate GPS technology allows for the simultaneous collection of geographic data with complex attributes and the creation or adjustment of geographic objects in the field. The use of handheld devices for address listing and boundary updates involves two related but distinct technologies: (1) a Computer-Assisted Personal Interviewing (CAPI) version of the address listing questionnaire and (2) a user-interface and processing engine for the manipulation of spatial data.

Custom applications are necessary for a handheld device to capture spatial data and demographic data concurrently.

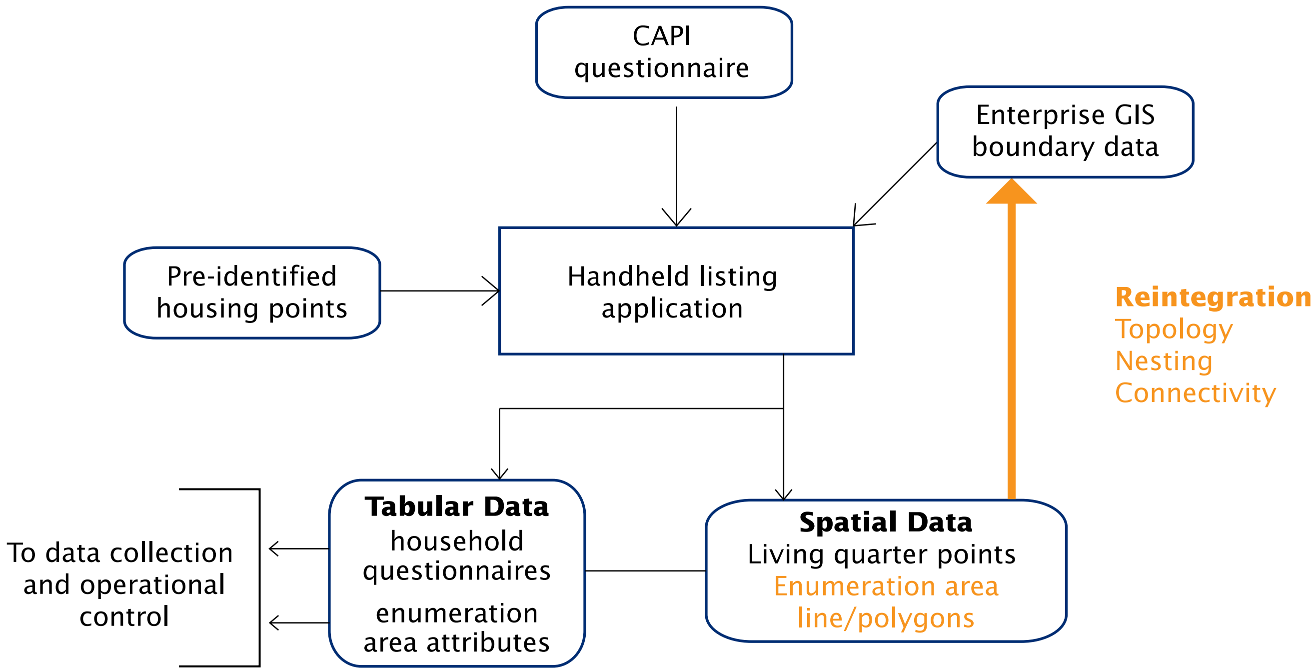
Custom tablet and smartphone geographic listing applications, when driven by geographically-enabled software, can access and manipulate all geographic primitives (points, lines, and polygons) on screen with a base map (satellite imagery or reference map) for field user guidance. Such applications can collect data to split building point datasets into individual living quarters, capture the number of housing units, and collect household-level data.

Point data are relatively easy to collect and manipulate on handheld devices. However, manipulation of boundaries in the field adds a substantial level of complexity to both the listing program and the workflow for reintegrating data collected in the field back into the geographic hierarchy. Reintegration of linear or polygon data manipulated in the field requires an advanced workflow ensuring that the topology of the nested boundary hierarchy is preserved. Furthermore, if multiple users are editing these data simultaneously, an enterprise GIS solution may need to be acquired (see next section) to effectively manage the flow of data. Figure 3 shows elements in orange that are required when a listing application includes a field update of statistical boundaries. A simpler solution for NSOs to consider is to use up-to-date high-resolution imagery for most enumeration area boundary updates.

Other important considerations when adopting handheld devices for census and survey operations are the locational

[2] The first commercially available GPS receiver was the Magellan Nav 1000, released May 1989.

Figure 3.
Flow of Data Between Listing Operation Technology



Source: U.S. Census Bureau.

accuracy of the device and the operating system. Tablets and smartphones are designed primarily as communication devices for everyday consumers, and their location accuracy can be low. NSOs should consult with technology vendors to document the accuracy of their devices and conduct field testing in different environments (such as cities, suburbs, and rural areas). NSOs should also assess their needs against the software available for different handheld device operating systems, such as Android, iOS, and Windows.

ENTERPRISE GIS AND SPATIAL DATA MANAGEMENT

Creating and maintaining an enterprise GIS for the geographic listing operation is encouraged early in census program planning. The "enterprise" refers to an office or organization, and an enterprise GIS ensures data flows easily between teams and individuals without compromising data quality or security.

An enterprise spatial database (or geodatabase) can store and manipulate spatial data and is typically administered through a relational database management system (RDBMS). This database could be accessible within a specific working group (such as the GIS staff) or it could be available throughout the NSO and possibly via a secure Internet portal, depending on the requirements of the listing operation.

Enterprise GIS ensures data flows easily between teams and individuals without compromising data quality or security.

NSOs will have to decide between either proprietary commercial software or free and open source software (FOSS) when implementing their enterprise GIS framework for the geographic listing operation. As shown in Figure 4, each option has strengths and weaknesses. NSOs should discuss these software considerations with their staff (if keeping development in-house) or chosen vendor (if outsourcing) early in census planning. Regardless, for both open source and proprietary software, NSOs must ensure the enterprise GIS can be maintained over the entire census or survey lifecycle and beyond.

Open source and proprietary software can be used for different components of the enterprise GIS. Many of these components are interoperable, meaning an NSO could use an open source solution for one component but a proprietary solution for another, depending on the workflow

Figure 4.
Key Distinctions Between Free and Open Source Software (FOSS) and Proprietary Software

	Free and Open Source Software	Proprietary Software
Licensing fees	None.	May have up-front fees and/or annual maintenance fees.
Source code	Complete access, providing greater customization options for software developers. Not all free software is open source.	Not open to the public and protected by copyright.
Ease-of-use	May be heavily reliant on command-line interface and user programming knowledge, requiring more expertise.	Typically a user-friendly graphical interface, requiring less expertise.
Technical support	Limited to online user community; dedicated support possibly available for purchase from private vendor.	Typically available directly from software publisher.

Source: U.S.Census Bureau.

requirements. Figure 5 shows examples of well-known open source and proprietary GIS software.

The solutions discussed above are typically hosted by NSOs with on-site servers or through a cloud service and require direct management by the NSO or a contracted vendor. However, an emerging set of alternative solutions are categorized in Figure 5 as GIS-as-a-Service. These solutions are a hybrid of database, server, and Web map and offer varying degrees of functionality. A GIS-as-a-Service solution can potentially reduce the human and physical capital required to host geospatial data and reduce operational costs. However, GIS-as-a-Service provides less control and customization than an on-site server solution and may be preferable for the dissemination phase of the census or survey lifecycle rather than the operational phase.

ASSESSING AND BUILDING GIS STAFF CAPACITY

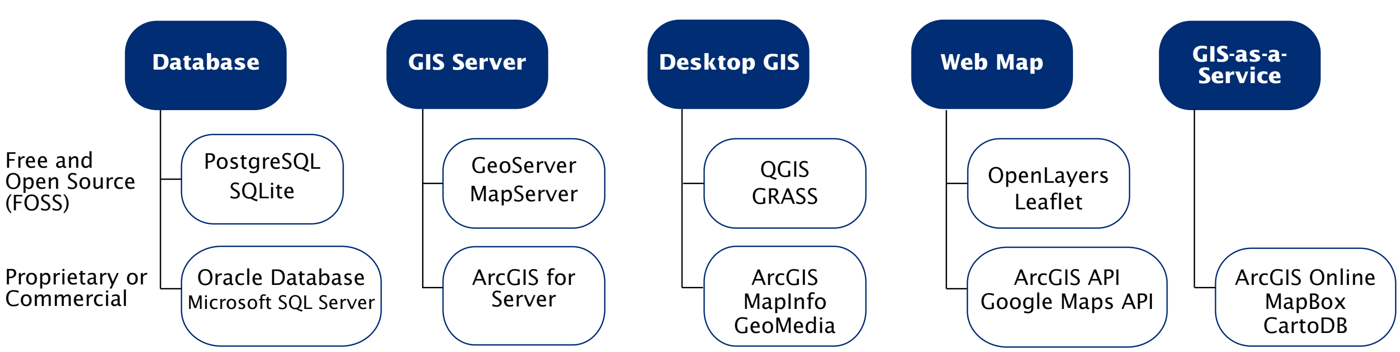
The emerging geospatial technologies discussed in this brief require a reconsideration of the technical skills held by GIS staff in a NSO.

Traditionally, geographers and cartographers had overlapping skillsets: geographers performed tasks such as devising classifications for geographic areas and analyzing spatial demographic trends, while cartographers produced maps used in field and dissemination materials. Increased use of desktop GIS has led to a convergence of these occupations as the time and effort needed to make maps has decreased.

However, the skills required to implement modern geospatial technology have evolved (Figure 6). Increasingly, staff must have the ability to automate GIS operations with scripting languages, design databases for storing geospatial data, or build interactive Web maps. Hiring a single individual with the full range of skills necessary to take advantage of these new technologies is difficult, but not all NSO GIS staff must become experts in all of these skills. Therefore, careful consideration should be given to the distribution

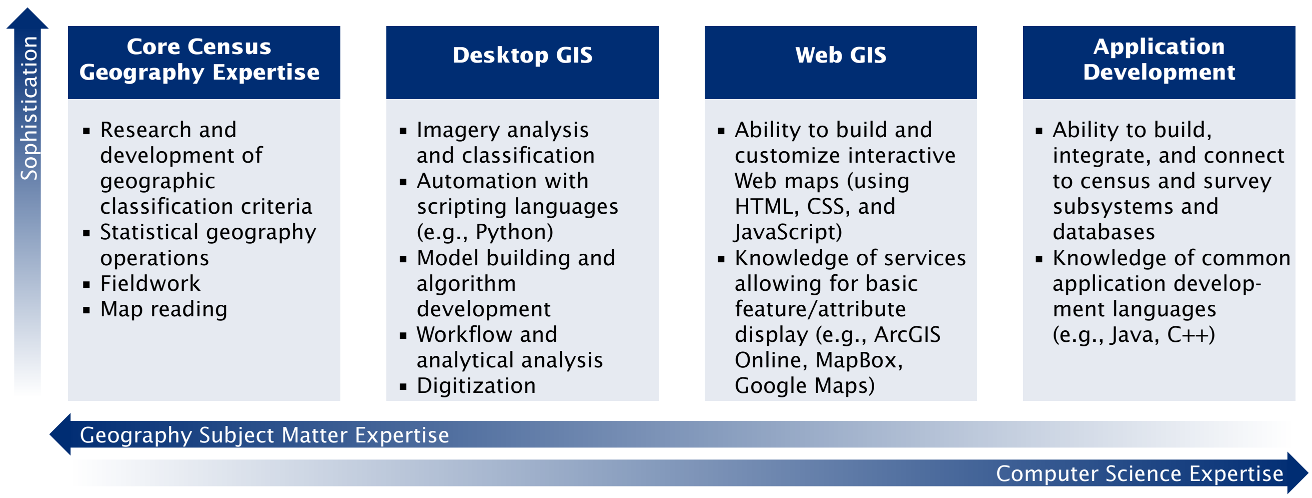
Successful GIS managers will develop areas with varied skills while encouraging cross-area understanding, cooperation, and collaboration.

Figure 5.
Examples of Enterprise GIS Software Components



Source: U.S. Census Bureau.

Figure 6.
Areas of Expertise Required for Successful Implementation of New GIS Technology



Source: U.S. Census Bureau.

of responsibilities among the GIS staff and to the type of expertise needed to fulfill a particular goal. Successful GIS managers will develop areas with varied skills while encouraging cross-area understanding, cooperation, and collaboration. Managers should also encourage their GIS staff to conduct independent problem solving and self-teaching, considering the rapidly evolving and highly technical nature of GIS.

OUTSOURCING

After assessing NSO GIS staff abilities, it may not be feasible to develop the in-house skills necessary to build, integrate, and deploy a complete geographic listing system. In this case, an NSO may outsource the development of such a system to a contracted private vendor. The primary goal of outsourcing should be to gain temporary access to skills otherwise not available within the NSO or to augment the amount of staff available with a certain set of skills. Consider these guidelines if outsourcing:

1. Do not cede full control of the system design and development to the contracted vendor. Responsibility for the ultimate success or failure of the listing operation must remain with the NSO, not the vendor. Therefore, the NSO must understand the technology solution being implemented and its risks.

Responsibility for the ultimate success or failure of the listing operation must remain with the NSO, not the vendor.

2. Make use of the institutional knowledge of GIS staff that have experience with statistical boundary update operations and address listing.

3. Document the workflow for updating enumeration maps prior to the use of new geospatial technologies and use these workflows to design the address listing system with the vendor.

4. Do not let technology drive the design of the listing system.

NSOs must take into consideration future maintenance, expandability, and staff skills when considering a vendor. Vendors often specialize in either proprietary or open source systems.

OTHER IMPORTANT CONSIDERATIONS

There are risks associated with adopting new technologies. Existing staff members may not welcome changes to their workflows and resist the adoption of new technologies. Additionally, new technologies require investments in data security and staff training to prevent the loss of sensitive individual data.

New technology should also be tested thoroughly before deployment. Such testing must be built into the project schedule and be provided adequate time to make improvements prior to implementation. If testing shows a software solution will not be ready in time for the operation, an alternative plan must be available to ensure successful completion of the operation.

GIS staff, GIS managers, and senior leadership all play critical roles and share accountability for ensuring the successful implementation of new technologies at NSOs.

REFERENCES

Jensen, John R., *Remote Sensing of the Environment: An Earth Resource Perspective,* Second Edition, Prentice Hall, Upper Saddle River, NJ, 2006.

______, *Introductory Digital Image Processing*, Third Edition, Prentice Hall, Upper Saddle River, NJ, 2005.

Longley, Paul A., Michael F. Goodchild, David J. Maguire, and David W. Rhind, *Geographic Information Systems and Science,* Third Edition, Wiley, Hoboken, NJ, 2011.

United Nations Statistics Division, *Principles and Recommendations for Population and Housing Censuses,* Revision 3, United Nations Publications, New York, 2015.

______, *Handbook on Geospatial Infrastructure in Support of Census Activities*, United Nations Publications, New York, 2009.