United Nations Regional Workshop on the Use of Electronic Data Collection Technologies in Population and Housing Censuses

29-31 January 2018
Cairo

Session 5

Use of Geospatial information in support of census operations

United Nations Statistics Division

Data collection with handheld electronic devices

- The Global Statistical Geospatial Framework
- Geocoding: Locating data to map
- Geographic databases
- Role of geospatial information in census operations
  - Pre-enumeration
  - During enumeration
  - Post-enumeration
- Planning considerations for developing a census geographic program

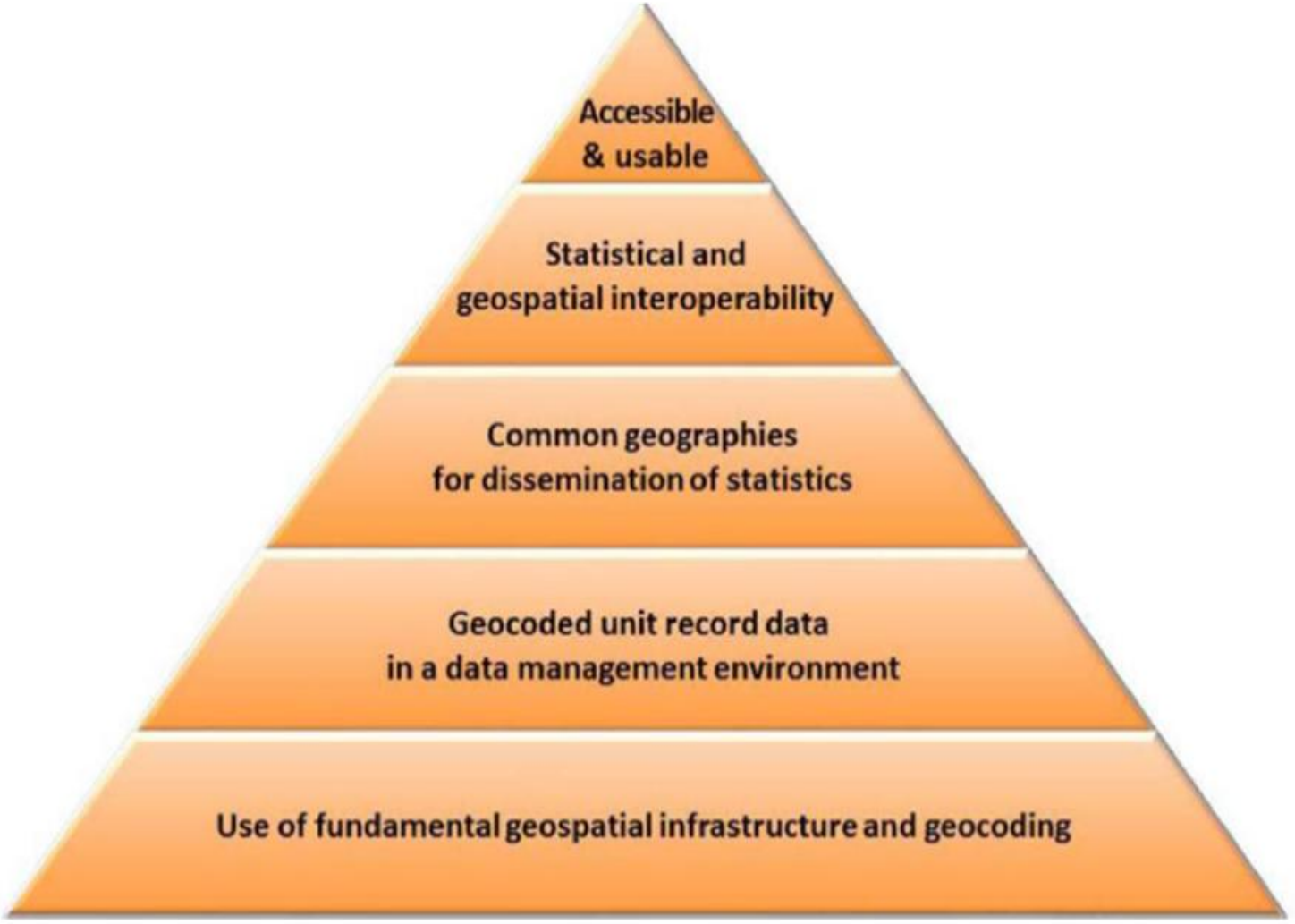
Introduction

- Census geography is essential to plan and manage fieldwork as well as to report results
- New geospatial capabilities have enabled NSOs to collect more accurate and timely information about their populations as a result of technological advances in:
  - Global Navigation Satellite Systems (GNSS)
  - Geographic information systems (GIS)
  - Availability of affordable aerial and satellite imagery
- Geospatial technology& tools have enabled more efficient production of both enumerator& thematic maps, improving the overall census data quality and supporting more efficient and impactful analysis of census data

The Global Statistical Geospatial Framework

- The GSGF framework consists of five guiding principles for integrating statistical and geospatial data
  - adopted in Augusts 2016 by the 6th Session of the UN Committee of Experts on Global Geospatial Information Management (Decision 6/107) – set up by UN Stat Comm and UN-GGIM
  - recognizes the 2030 Agenda & the 2020 WPHCP as important drivers for the integration of geospatial and statistical information in support of evidence-based decision-making

The Global Statistical Geospatial Framework



The Global Statistical Geospatial Framework

- The Five Principles of the Global Geospatial Statistical Framework are:
  1. Use of fundamental geospatial infrastructure and geocoding
  2. Geocoded unit record data in a data management environment
  3. Common geographies for dissemination of statistics
  4. Interoperable data and metadata standards
  5. Accessible and usable geospatially enabled statistics

Geocoding: Locating data to map

- Principle 1 of the GSGF considers geocoding as a fundamental principle for integrating geospatial and statistical info.
- Geocoding is the process of assigning geographic coordinates (e.g. X-Y values, latitudes& longitudes) to specific locations (such as street addresses & place names) so that they can be placed as points on the earth's surface (&therefore on a map)
- This provides for a common and consistent approach to establishing the location and a “Geocode” for each unit in dataset, whether it is an enumeration area, census block, a building, or parcel/property unit

Geocoding: Locating data to map

- Geocoding is indispensable for locating buildings and geographic features, especially in areas lacking street infrastructure or a reliable address system
- Geocoding is critical for the transition into a fully GIS based approach to census mapping
- Geocoding is vital for producing high quality maps and performing census tasks such as:
  - modelling and delineating administrative and EA boundaries
  - point locations of census units: housing units& collective living qrt
  - locating other relevant geographic features such as roads, rivers and landmarks
  - supporting the dissemination, aggregation and disaggregation of data; aggregate data into new/customized units of analysis

Geographic databases

- Principle 2 of the GSGF recommends storage of geographic information and data into databases
- Geodatabases allow for data to be (re)utilised for future purposes and needs, including for creating new geographies in a time-efficient and geographically consistent manner
- Geodatabases are key GIS tools for maximizing the use of geographically referenced data and forms the basis for all queries, analysis and decision-making within a GIS
- The development of the digital census database is usually based on two data sources:
  - the conversion and integration of existing geographic data/maps that may be in hardcopy or digital form and
  - the collection of additional data using fieldwork with GNSS receiver or digitization of satellite and aerial imagery and other raster files

Geographic databases

- A comprehensive census database usually consists the following elements:
  - *spatial boundary* database, consisting of area features (polygons) that represent the census units (e.g. EAs and administrative/statistical divisions)
  - *geographic attributes table*, a database file linked internally to the spatial database that contains one record for each polygon. This table contains the unique identifier for each census unit and possibly some additional static or unchanging variables, such as the unit's area in km2
  - *census data tables* containing non-spatial attributes, i.e., the census indicators for the spatial census units. Each of these files must contain the unique identifier of the census unit that provides the link to the corresponding polygon attribute table records. There will be one record for each census unit
  - *other vector (point or area) features,* such as building/housing unit points, landmarks, roads, waterways, schools, health facilities or other buildings may be useful for orienting fieldworkers during the enumeration. Such features are usually recorded during preliminary field-canvassing or house-listing

Role of geospatial information in census operations

- Role of maps in the census process has expanded
- Traditionally: to support enumeration, operations management and dissemination
- Today, through use of modern GIS system, geospatial information helps to:
  - improve efficiencies and accuracy of the overall census project and products
  - optimize enumeration areas, workforce assignments and field offices
  - save cost (e.g. labour and transport) in census operations

Pre-enumeration

- Create/update census geographic database
- Create/update Base Maps
- Create Map Services
- Create or update enumeration areas (EAs)
  - Produce digital EA maps for fieldwork and operations
  - Use of remote sensed data (images from satellite and aerial imagery) in EA creation
  - Integrate fieldwork using remote sensed data
- Validate EA's
- Ensure complete and balanced coverage (no omissions or duplications) through GIS analysis
- Facilitate efficient census operations (eg. determine the most efficient placement of field offices and allocation of field workers) through the application of GIS analysis

During enumeration: Workforce management

- Assign work to individual enumerators, monitor their activities and provide necessary assistance
- Improve time and resource management
  - reduce labour costs, optimize travel time, travel costs
  - improve the overall quality of the census
- Create efficiencies and provide transparency across the workforce
  - forecasting of workload and required staff
  - management of working times
  - suggest best possible scenarios for completion of work
  - deploy technical or specialized skills and equipment

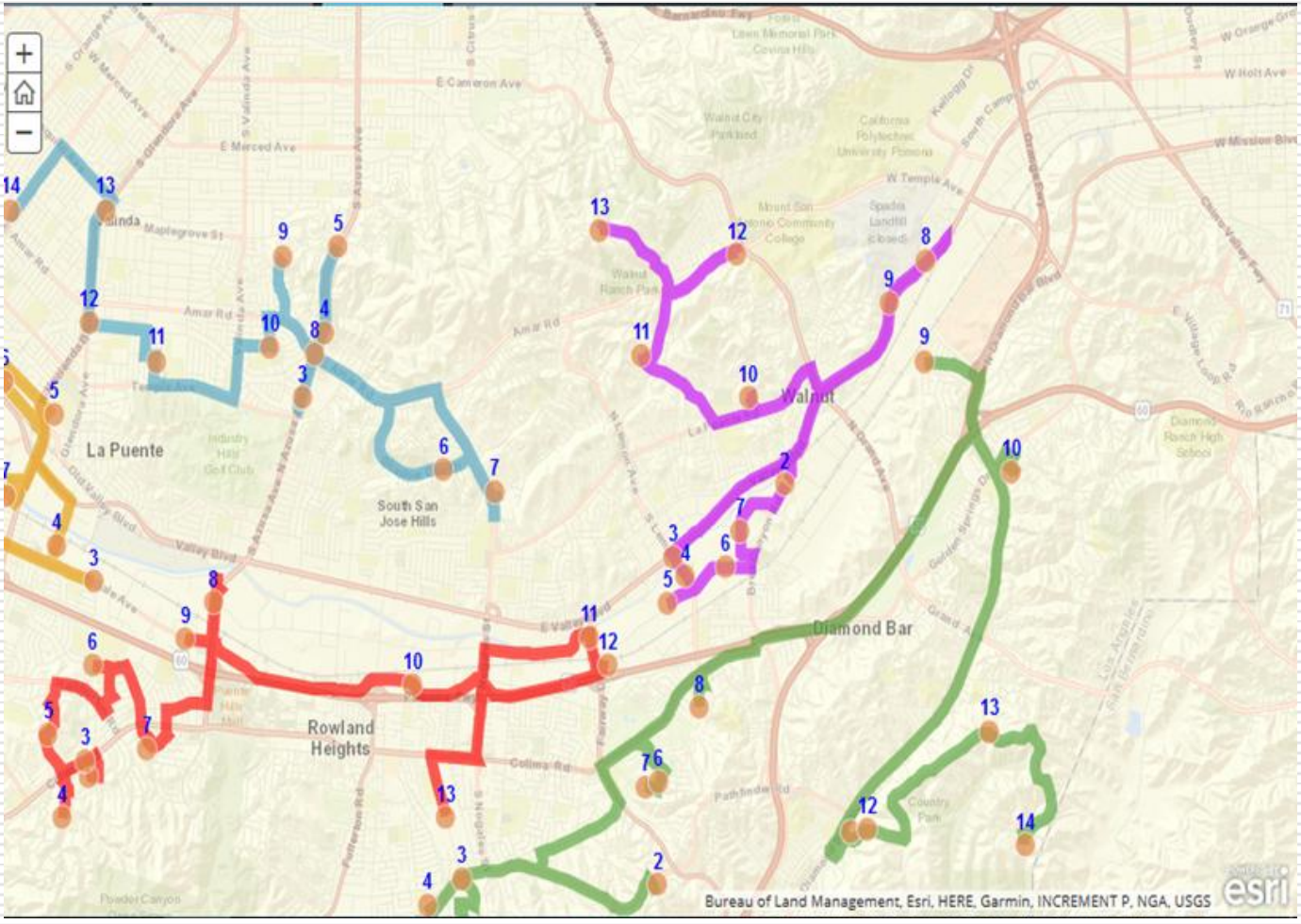
During enumeration: Workforce management

- GIS solutions allow for complex scheduling analysis and assignments
  - optimize routes by determining quickest/shortest/most cost-effective route through EA
  - find closest facilities in order to estimate cost of traveling between locations
  - pre-plan routes, saving time, labour and fuel for transport

During enumeration: Workforce management

- **Example of daily optimized routes for one enumerator**

  - the routes for each day are depicted in different colours and include the optimal stop sequence number for each site visit



During enumeration: Monitoring and operation management

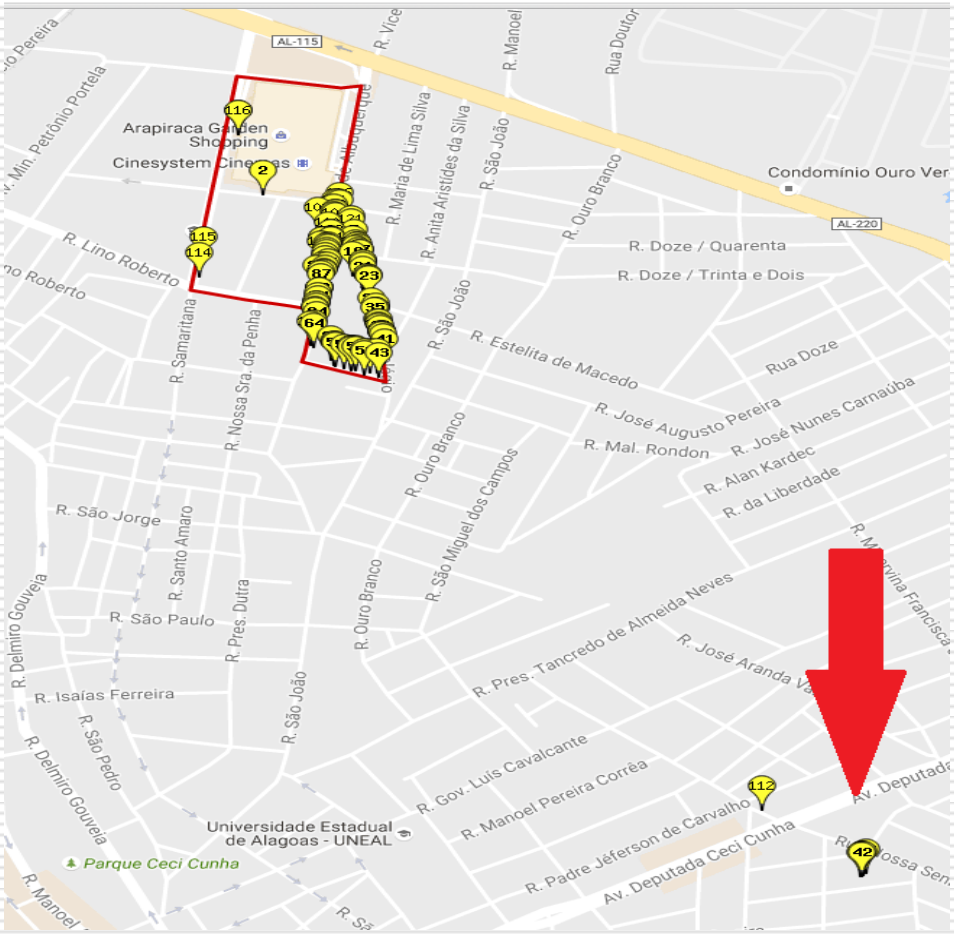
- Support planning and control tasks by supervisors
  - make assignments, reallocate tasks, with the aim of ensuring a smooth and timely completion of the enumeration
  - real-time transmission of geo-data of enumerator route and household locations enable the supervisor to monitor the progress of the census enumeration locally
  - facilitate better management of the enumerator's workload by checking the effective progress, comparing the actual completions rates compared with those expected
  - identify problem areas and implement remedial action quickly
  -

During enumeration: Monitoring and operation management

- Brings together a common view of the progress of the enumeration through GIS operational dashboards `
  - allow for monitoring of real-time data feeds for day-to-day operations
  - monitor progress at any level of geographic aggregation
  - deliver information products needed throughout the operations processes including; maps, charts, and operational dashboards that provide an overview of the project status.
  - Provide updates on progress to decision makers

During enumeration: Monitoring and operation management

- Example of an attempt to complete a questionnaire at a location beyond the boundary of an enumerator's EA



During enumeration: Support the work of enumerators

- Puts GIS mapping in the hands of the field staff which improves the accuracy of collected census data
- Ensure enumerators can easily identify their assigned geographic areas, follow a pre-determined path, and know their exact location in real time
- Makes it easy for enumerators to organize their daily workloads by determining the optimum route to each household, report progress
- Support an alert function on the mobile applications to ensure enumerators' safety through notification of supervisor in any emergency situations

During enumeration: Updating and correction of EA maps during enumeration

- Census maps are usually prepared several months, or even years, ahead of the actual enumeration
  - new constructions and infrastructure developments may not be shown in the EA maps
  - census maps may contain errors that may lead to either under- or over-coverage
- It is usual practice to verify EAs and units to be covered just prior to the actual enumeration.
  - this round of validation of EAs is aimed at verifying the existence of buildings and supplementing the address list with missing address points
  - this data will allow the statistical agency to correct and update the digital maps in good time for the enumeration itself

Role of geospatial information in census operations: Post-enumeration

- GIS & interactive mapping make it easier to present, analyse and disseminate census results at various levels of geo. disaggregation
- Provide a powerful means for visualizing the results of a census & for identifying patterns among demographic and social indicators
- Allow data at census unit level to be aggregated to new units of analysis (such as climatic zones or ethnic regions)
- Link information from many different subject areas, leading to a much wider use of statistical information
- Make accessible vast amounts of spatial information to users through the Internet

Planning considerations for developing a census geographic program

- Key considerations include:
  - Needs assessment
  - Inventory of existing data sources
  - Collection of new geographic data
  - Geospatial data management system
  - Identification of geographic products and services
  - Staff skills and capacity
  - Outsourcing

Planning - Needs assessment

- Understanding user needs and requirements is the first step in planning for a census geography program
- Needs assessment should be based on consultations with stakeholders and users
- Consultation should be on both geographic content and products
- It is important to reconcile user expectations with what is feasible given available resources and constraints – trade-off between cost, time and quality

Planning - Inventory of existing data sources

- Identification of maps and other geographic information (both digital and hard-copy)
  - reduce need to collect additional data using fieldwork &/or imagery
  - save time by keeping field-checking to a minimum, and focusing on rapidly changing areas
- Cooperation with National Mapping Agencies, relevant government agencies, private sector
- Leverage resources of National Spatial Data Infrastructure (NSDI) institutions

Planning - Collection of new geographic data

- Two data-collection approaches:
  - collection of data using satellite or aerial imagery
  - collection of X-Y coordinate data using fieldwork
- Analysis of imagery acquired by satellites or aircraft
  - should be followed up with field veri-fication using handheld GNSS technology
- Field X-Y coordinate data collection using handheld devices
  - accurate capture of point locations
  - collecting boundaries challenging
  - if multiple users, enterprise GIS solution may be needed

Planning - Geospatial data management system

- GIS should ensure data flows b/n teams & individuals (in field & at headquarters)
- Spatial database (or geodatabase)
- GIS software - the processing engine, vital component of an operational GIS
  - commercial off-the-shelf software (COTS)
  - free and open source software (FOSS)
  - alternative solutions - GIS-as-a-Service

Planning - Identification of geographic products and services

- Identify the types of maps and map services needed in all phases of the census operation
- Take into consideration the available financial and human resources and other various factors
- Range of geographic outputs and products include:
  - digital maps for EA & dissemination areas;
  - geographic boundary files in digital format for all statistical reporting units
  - listings of all statistical and administrative reporting units, including geographic X-Y coordinates;

Planning - Identification of geographic products and services

- Range of geographic outputs and products include (cont'd):
  - geographic equivalency files that indicate how current reporting units relate to those used in previous censuses
  - vector layers containing feature data, such as buildings, landmarks, roads, schools, hospitals, etc
  - street index listings for all major urban areas;
  - centroid files that provide a representative geographic point reference for each reporting unit;
- New products: map services available on the web
- Proper documentation, including coding and metadata

Planning - Staff skills and capacity

- Technical capacity and skills held by GIS staff critical
- Increased use of GIS packages requires considerable training
- Developing GIS capacity may entail reorganization and expanding the existing “cartographic unit”
- Skills needed include:
  - planning and project management;
  - systems administration.
  - geographical data conversion;
  - map scanning and digitizing;
  - cartographic design;
  - field work with GNSS devices

Planning - Outsourcing

- If in-house skills needed to build, integrate & deploy a geogrphc listing system is lacking, outsourcing may be a solution
- 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
- The following guidelines should be con-sidered if outsourcing:
  - Do not cede full control of the system design and devel-opment to the con-tracted vendor
  - Respon-sibility for the ultimate success or failure of the geographic operation must remain with the NSO, not the vendor
  - Document the workflow for updating enumeration maps prior to the use of new geospatial technologies
  - Do not let technology drive the design of the geographic system
  - NSOs must take into consideration future maintenance, expandability, and staff skills when considering a vendor