الورشة التدريبية عن بعد حول تعزيز القدرة الجغرافية المكانية في تعدادات السكان والمساكن للبلدان العربية

20-23 يوليو / تموز 2020

منصة المعهد العربي للتدريب والبحوث الإحصائية
Use of Mobile Devices: Integrated Field Data Collection

Amor Laaribi, PhD
UN Retiree, Former UN-GGIM Coordinator at UNSD/DESA
International Freelance Consultant in Geospatial Information
Outline

- Challenges with Paper-based Data Collection/Processing
- Moving to Handheld-based Data Collection/Processing
  - An Integrated Approach
  - Key Features
  - Advantages & Disadvantages
- Implementation/Organizational Aspects
- Challenges
- Recommendations/Conclusions
Paper-based Data Collection

- Paper-and-Pencil Interviewing (PAPI)
- Self-administrated Questionnaire

- Handheld devices
- Internet
Paper-based Approach

Data Collection

Data Entry

Data Coding/Editing

Manual

- Key entry/computer-assisted
  - OMR/OCR/ICR

Manual Computer-assisted
Challenges with Paper-based Data Collection

- Inefficient information gathering
  - Slow and time-consuming
  - Re-keying information is inefficient and increases chance of error
  - Paper-based forms increase chance of error
  - Submitting multiple forms for a single process
  - Cannot capture value-added data (e.g. geographic data)

- Data Integrity and Authenticity:
  - No way to check that the data entered is correct
  - No cross checking with third-party data sources
  - With this approach, data accuracy and timeliness are big concerns!
Traditional paper-based methods of census data collection have proven to be tedious, time consuming, costly and often prone to errors.

To overcome these problems, Computer Assisted Personal Interviewing (CAPI) methods are increasingly replacing pen-and-paper methods as a viable alternative for census data collection.

The technological developments in mobile computing, storage, higher connectivity and network coverage have made the CAPI approach possible for many.
As increasingly powerful hand-held devices are now an affordable and realistic option, the CAPI approach is being more commonly used.

Realizing the advantages of direct digital data entry into a computer software application at the point of data collection through the use mobile technology for data collection and statistical production, the United Nations has recommended its use in the 2020 Round of Censuses in recognition of its importance and usefulness for many countries.
In a survey on national experiences for PHCs of the 2010 Round conducted by UNSD in 2012:

- Most countries reported that the implementation of new technologies was the most successful aspect of the census-taking conducted for the 2010 Round of Censuses, with 56% of the countries responding to the questionnaire.

- Innovative technologies included the use of mobile technology (hand-held devices) and Internet for data collection; geospatial technologies for mapping, and other web-based and mobile telephony applications for dissemination.
“CAPI is the face-to-face interviewing mode in which a computer displays the questions onscreen, the interviewer reads then to the respondent, and enters the respondent’s answers into the computer”.

Hand-held devices, such as such as Personal Digital Assistants (PDAs), Handheld computers, Smart phones and particularly Tablets are emerging as powerful tools, improving the quality of the data collected, the timeliness of census releases and reduced costs of field operations.
Trend in Data Collection: Moving to Handheld Devices (Tablet computer)

- Personal Digital Assistant:
  - “A personal digital assistant (PDA) is a handheld computer, allowing data to be captured and stored electronically.”

- PDA vs Tablet computer:
  - **Handheld computer/Tablet** offers more features than a PDA (larger than a PDA; smaller than a Laptop!)
  - Used interchangeably (Mobile Device)

- Integrated field platforms with:
  - GPS; cellular; camera; cellular network connection for voice, SMS, and Internet data communication, and OS applications

- Increases Productivity “on the move”
  - Data can be collected in digital format in the field
  - Data can be directly uploaded into the server
Electronic questionnaire - Contents of the census form are stored onto the Tablet so that the questions appear sequentially on the screen

Data are entered into a hand-held computer instead of onto a paper census form, allowing:

- Immediate evaluation at the moment of data collection, allowing the correction of information at the moment of the interview;

- The filling out of all the compulsory questions, avoiding the lack of answers due to forgetfulness or mistake by the enumerator;

- Optimization of the filling out of data through automatic skips in the questionnaire, avoiding covering several items about which, sometimes, there would be no reply; which could;

- Optimize time used by the enumerator and the head of household.
Besides providing:
- Mobile/Electronic forms
- Inbuilt validations for data
- Highly Secure: Biometrics and Wireless

Data are then electronically transmitted to an Central Data Center for further processing:
- Offline or online depending upon the availability of connectivity options: WiFi, CDMA / GSM Radio
  - For example, if work is done in a remote area with no connectivity, the data would be stored in the device itself, but when we reach an area with connectivity the data would be automatically transferred to the Data Center

Other characteristics:
- Ease of use
- Multi-lingual capability
Handheld devices with In-built GPS

The device can be enabled with GPS to:

- Access to coordinates of the units visited during data collection;

- Use of coordinates obtained during data collection to track the location of the place from where the data was entered which would allow the department to check cases of fraudulent data entry; and

- Tracking could be undertaken to assist the enumerator in understanding their current location and also capture the geographical location of where the census data was captured.
Data security

Data at rest—stored on the tablet PC storage media
- Login password protection (OS and enumeration system)
- Basic user privileges for census officers
  - Use of the enumeration software system
  - Data synchronization
- File system level encryption of the local database files
- Census officers do not have access to the decryption key
- USB flash drives disabled

Data in transit—during synchronization
- From tablet PC to work branch
  - Wi-Fi 802.11n WPA2
  - IPsec/L2TP VPN using digital certificates
- From work branch to HQ
  - Private network infrastructure
  - IPsec VPN tunneling
The suitable for census data collection would range from 7 to 10 inches, due to the optimized size and weight, brightness, and for being held with one hand, all this suitable for field work.

And rugged tablets needed for the tough field environment, and to avoid breaking and dissuade theft.
<table>
<thead>
<tr>
<th>Feature</th>
<th>Details</th>
</tr>
</thead>
</table>
| Affordability: Price of the device          | - License/subscription cost or purchase fees  
- Maintenance fees  
- Annual/monthly renewal fees |
| Interface                                   | - User interfaces, color, resolution, keyboard |
| Battery life and management:                | - Length of battery life |
| OS updates:                                 | - How often is OS updated with bug releases and features |
| Customizability:                            | - Can the device be customized including shortcuts, color schemes, keyboards, etc. |
| Storage/Memory                              | - Storage capacity |
| Peripheral (or APP) support                | - The ecosystem of peripherals or apps (GPS, Camera, etc) |
Field Operations Management

- Using new technologies including GIS is the capabilities they provide the NSO to streamline and automate field operations and thus improve their management and the quality of the census itself.

- Field operations require efficient communication between massive number of fieldworkers, field supervisors and census managers, as well as constant monitoring of a large variety of logistical items.

- Mobile technology provides real time bi-directional flow of information between census managers in the back-office and enumerators in the field.

- Bi-directional workflows allow the census managers to be informed of the progress of the data collection operations: while managers provide the enumerators with updates, including which households need to be followed-up

Source: Esri Book; GIS and the 2020 Census
Other Benefits of Geospatial Technologies to the field Operations

- With GPS enabled handheld devices, census managers can locate and track the location of the enumerator (in real time or near real time) or identify areas where there are gaps where enumeration is falling behind, or not meeting quality standards.

- This triggers urgent attention to be given to the impacted EAs for appropriate decisions in order to remedy the situation.

- In this manner GIS systems are providing for better engagement across all levels of management and field workers, providing needed transparency.
Operational dashboards can be used to provide a view of geographic information that helps you monitor events or activities. Dashboards are designed to display multiple visualizations that work together on a single screen. From a dynamic dashboard, supervisors can view the activities and key performance indicators most vital to meeting objectives.
Field Operations workflows

- **Understand & Communicate**: Plan, organize, prioritize and assign work.
- **Coordinate**: Receive assignments, communicate status and coordinate with others.
- **Monitor**: Ensure data quality and integrity, make better decisions.
- **Navigate**: Get where you need to be, optimize your travel routes, use your own data.
- **Capture**: Conduct surveys, inspections, collect new information and document status.
- **ArcGIS**: Monitor the progress and effectiveness of your field operations.

Source: Esri Book; GIS and the 2020 Census
Summary of the Key Features

- Real-time task management from office
- Manage field time, expenses, etc.
- Data synchronization and management
- Secure applications, procedures and protocols
- Ability to encrypt data for secure transmission
- Integrated field platforms with GPS, cellular, camera and OS applications
- Geographical and map processing
Advantages vs Challenges of Use of Handheld Devices

**Advantages**
- Instant data capturing at the point of collection, reducing manual input errors
- Immediate data validation, reducing re-verifications at later stage
- Time effective with real time logical validation rules, reducing logical errors
- Faster processing of census information leading to timely availability of results
- Integration of geospatial data; other value-added data
- The use of Handheld devices during enumeration raises the profile of digital communication with the general population especially with those who see this type of technology for the first time.

**Challenges**
- Setting up of process may take a long time as it requires extensive testing
- Requires that enumerators have ability to use the device which may require administering a test
- Requires intensive training of enumerators on use of device (training is more complicated)
- Need to recharge the battery which could run out during enumeration
- Possibility of equipment failure, theft, etc. In the unlikely event that the unit is lost or stolen, precautions should be in place to transfer data from the device and remotely wipe the device.
Using handheld device would significantly automate the whole process of data collection by having a centralized Data Centre (and or a Regional Data Centre) where the data entered into the device would be collated automatically. The Data Centre would also enable the supervisors of the census collection process to make real time checks into the data collected to ratify that the data collected is relevant and correct.
Brazil: an example of massive use of PDAs

Case of Brazil:
- In-house devices
- First use in Agriculture Census in 2007
- Then, massive use in 2010 Population and Housing Census
  - 240,000 devices
- Deployed/borrowed to other countries: Cape Verde, Senegal, etc.
Example of Mobile Application - Poland

**Smartphone**
- Mobile app
- VPN network

**Census module**
- Tasks
- Questionnaire

**Map module**
- GPS
- Ortho-imagery
- Cadastral data
- Tasks

**Communication module**
- Phone (to dispatcher only)
- Messaging
- Alarm procedure

Central Statistical Office Poland
Mobile application

Central Statistical Office Poland
Mobile application – map module
Integration of CAPI App and GIS

- CAPI and GIS-based EA maps
- Satellite imagery as backdrop

Integration may need programming

Examples:

- Survey Solutions (World Bank)
- Survey123 and Collector

Suvery123 for ArcGIS on a handheld device
Need to Build Partnership with:
- Application development partner
  (IT Cie with expertise in mobile forms and hosting data centers)
- Device Manufacturer
  (To provide the devices as per specification)
- Connectivity provider
  (To provide connectivity for the device so that the data can be transferred seamlessly to the data center)
- Capacity building supporter
  (Training on using not only the forms and the entire process of data collection but also on the basics of the device and what to do for trouble shooting).

Nodal Agency:
- Operationalize the whole process
Significant costs (Time and money)

- For implementation of both Tablet or hand-held GPS, the following needs to be considered:
  - Purchase of hardware (PDA ~ $100-300/Each)
  - Technical Training for enumerators
  - Cost of Pre/Post Census mapping exercise
  - Software development
  - Logistical costs

- Considering/Planning for the deployment of the massive number of devices
Lessons Learned-Good Practices

- Learn about some practices/experiences: the qualitative as well as quantitative benefits of handheld devices have been proven in field in many countries (Australia, Brazil, Canada, Malaysia, New Zealand, Oman, Jordan, Cape Verde, etc…)

- Various Options are available for selecting handheld devices

- Clear identification of objective is required for selecting best device

- Important to have extensive training prior to deployment

- Build a solid partnership: Integrator

- Post implementation support – technical as well as hardware support ensures project success
CAWI on-line Internet questionnaire is not a simple replication of the paper form.

Adjustments are likely to be made to accommodate the respondent with an easy to access user-friendly interface.

CAWI data collection method relies on self-enumeration, and like other modes must ensure that every household and individual is counted once and only once.

GIS functionality can be included in the web based form by way of a map presented to the user allowing them to indicate their place of residence. This information can be captured and stored along with the form providing the location component of the survey for validation against pre-assigned codes or other information.

A key factor in managing data collection requires the provision of each respondent with a unique code.

This code may be linked to a geographic location (e.g. the “Census-assigned unique address identifier” to be used at the 2020 US Census).

Since the CAWI collected data has to be incorporated into the database with other modes of collected data (CAPI/PAPI), census data collection methodology requires careful thought as does data management of the other streams of data, when the census is multi-modal.

Users demand increasing precision:
Thus, a shift from polygons (administrative and statistical units) to smaller areas and until points of reference (housing units, dwellings, parcel units, buildings, landmarks, etc.).

Point of relevance
- Output areas, such as enumeration areas, districts or other small areas, do not meet the user demands on territorial flexibility.

- Point-based information facilitates the convergence of information from multiple sources for a particular location.

- Producing single coordinate points as the spatial key-element will be able to deliver aggregations at any spatial unit, and create the population-grid at needed scales.

- Point-based data related to statistical units can include: persons, housing units, dwellings, parcels/units of land, buildings, landmarks, businesses, etc.
A point-based location of people and dwellings increases the spatial relevance of statistical information, and becomes crucially important in crisis management.

“Understanding places and the "where?" of things is vitally important in disaster response, as in the case of flooding and other natural disasters such as storms and fires.

The point-based approach is increasingly used through the direct capture with GPS of the coordinates (latitude and longitude) of point-based features such as dwellings or other features of interest.

Other technologies are supporting point-based data collection: such as mobile technology, high-resolution imagery, UAV, etc.
Mobile devices for data collection/dissemination is being recognized as an option worth considering in the next round of censuses (some developing countries, particularly, find it an appropriate solution to the challenges faced with the paper-based approach). The United Nations has recommended its use in the 2020 Round of Censuses in recognition of its importance and usefulness for many countries.

As increasingly powerful hand-held devices are now an affordable and realistic option, the CAPI approach is being more commonly used.

Still, there is necessity of early planning and through preparation for the adoption of mobile devices solution, including prototyping, use in small survey projects, and pilot exercises in order to validate all the stages and anticipate alternative solutions – a gradual approach.
As for any new technology-based approach, need to seek partnership/outsourcing

Particular consideration for the “integration aspects” and training of enumerators

Part of a multi-modal approach

The use of hand-held devices equipped with GPS can especially contribute to the geocoding of schools, hospitals and other important reference elements, as a side benefit of the census operations. It is recommended to use mobile technology, ‘GPS’, satellite imagery, and UAV, as they are facilitating data collection at the individual level, provided we take into account the privacy/confidentiality issues.
 شكرا جزيلا

THANK YOU !!

Email: amorlaaribi7@gmail.com