(Excerpts from document: Innovations and impacts on the Brazilian statistical and geographic information systems)

PDAs and electronic questionnaires

The use of mobile technology contributed to make the process more practical, light, effective, fast and safe. From this moment on, the questionnaires are presented virtually on the screen of the portable PDA equipment, with the answers written on the screen itself, or selected from different options. The use of PDAs allows these answers to be analyzed and checked during the interviews, to be stored and transmitted to subsequent centralized processing.

The digital equipement had been previously used in other IBGE surveys. However, it was the first time the IBGE used the PDA in a census operation. The PDA, being widely employed, brought several advantages to data collection, among which can be mentioned:

- 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 to cover several items about which, sometimes, there would be no reply; that can optimize time use by the enumerator and the informant; and
- the non-necessity of the transportation of a big amount of paper questionnaires and of the handling of these same questionnaires in data capture centers, achieving information precision and better processing time.

Besides all these advantages, handheld computers, equipped with a receiver for GPS signals allowed the georeferencing of all the units visited in the rural areas, as well as the monitoring of the geographic coverage.

GPS georeferencing in data collection

The use of GPS equipment and PDAs represented, undoubtedly, one of the most important innovations in the 2007 Censuses. By allowing the access to quality coordinates adequate to a wide range of uses, the introduction of GPS created the necessary conditions for the development of a big number of innovative applications. As a result of impressive technological development, this equipment can, in a short period of time and at low cost, provide an operator who has little technical knowledge with precise coordinates, in a type of operation which used to involve complex calculations and procedures.

Four major processes have been developed for the use of GPS, as shown below:

- access to coordinates of spots in the urban census track perimeter during the pre-collection operation;
- access to coordinates of the units visited during data collection in rural census tracks;
- use of coordinates obtained during data collection in distance-supervision procedures; and
- use in the field of the coordinates obtained for positioning and guidance of the enumerator.

The collection of coordinates in the rural area was done in each existing unit, whereas in the urban area, it was limited to obtain the coordinates of spots in the perimeter of the census tracks. The choice was not to collect household coordinates in urban areas, due to constraints in the use

of GPS technology, especially in terms of what concerns interruption of the signs sent by satellites.
Follows some more information about each one of the uses previously mentioned:
Pre-collection was one of the several preparatory activities in the Censuses. It was aimed at the recognition of the census track by the supervisor, before the beginning of data collection, in order to provide support to the enumerator. In spite of the effort for the improvement of cartographic information, the Territorial Base was not suffi ciently updated in the years preceding the 2007 Censuses. This way, by requiring that the supervisor go over the census track under his responsibility on the days preceding the collection, there was also a last opportunity for update and identification of inconsistencies, especially in the boundaries which defined the census track limits.

With this objective, IBGE has conducted this pre-collection operation since the 2000 Census. In order to guarantee that, in a short period of time, over 10,000 supervisors cover their areas, each supervisor was instructed to capture; obtain coordinates from specific determining positions in each census track, by means of GPS. By requesting from the supervisors the obtaining of point coordinates limiting each census track with GPS, the monitoring of the work can be done in a simple way. Once getting the coordinates requires the movement over the census track, the data obtained serve as proof of the presence of the supervisor in the area indicated. According to the instructions received, the first spot obtained by the supervisor is associated to the initial spot of the census track defined in a document which presents the description of the census track, and can be identified in the map. This way, by marking the coordinate on the census track map, it is possible to confirm the completion of the operation.

During collection itself, coordinates from all the households and agricultural holdings in rural areas were obtained. Considering the difficulties of building a list of addresses in the rural areas, similar to that of urban areas, these coordinates functioned as an alternative address, making it easier to return to the unit georeferenced and registered. It also made possible to visualize the progress of collection, in the map of each census track.

Although coordinates are easily obtained, some specific situations may turn this process difficult or even impossible. Units located in deep valleys, or in areas with dense vegetation may turn the identification of coordinates impossible. In order to inform the enumerator about the possible loss of quality of the coordinate obtained, the system displayed a quality indicator which prompted the enumerator into looking for a better position. In case the difficulty remained, the enumerator should register this problem through an option available in the system. Considering agricultural holdings, the coordinates were supposed to be collected, preferably, in the place itself, or at the entrance of it.

The use of GPS also made it possible to monitor the evolution of collection by means of the visualization of coordinates obtained, superposed on orbital images, especially in more remote rural areas. The comparison of the coordinates obtained with signs of residential or agricultural occupation detected in the image (Figure 35) represented a real effective and low-cost supervision mechanism in rural areas. Comparison also made it possible to observe, in the office and after collection, inaccuracies due to census track invasion and their correct spatial distribution.

The availability of a GPS receiver in the PDA helped IBGE develop a software to better guide the movement of the enumerator in some rural areas, especially in the North and Central West Areas. In order to give support to the 2007 Census operations, IBGE used the geoprocessing software Geopad and customized it to fulfill the demands of field work. The software was then called IBGE Mobile GIS, and it allowed the enumerator to identify their position in the rural area, their working census track and the distance covered in the field.

Being installed in the PDAs, the software included spatial information, such as maps in shapefile format and orbital images in JPG format, georeferenced in a simple and dynamic way. It worked through basic commands such as move, zoom in and zoom out, in the location of the census track, in the identification of its boundaries the access ways and significant elements in the territory.

New functionalities are being incorporated to the program for use in future surveys, especially in the 2010 Census.

Data capture and transmission systems

These systems allow IBGE personnel to detect inconsistent data at the moment of their input in the system, preventing, this way, the return to the place of interview in order to validate incorrect information. The system also allows the transmission of information from the PDA directly to the IBGE database, without the need of the stages of data digitizing or scanning, which were necessary previously.

IBGE equipped 530 agencies which are permanently distributed over the Brazilian main municipalities and about 574 customer service units with computers and broadband internet access. Computerized Collecting Stations were created and, in these places, enumerators were able to connect their PDAs to microcomputers, using local wireless communication device – Bluetooth, more specifi cally – in order to transmit data from the PDAs to census computers located in Rio de Janeiro.

Taking into consideration the communication resources existing in all the Brazilian areas, the following communication system was chosen for these stations: about 400 of them used ADSL technology (a name given to public telephony commuted) and 700 stations had communication technology by satellite, when there weren't land means of transportation, especially in the North Region of the Country.

The connections were made with IBGE Networks, using the Internet, through an exclusive software for transmission of cryptographic data (Virtual Private Network - VPN), which guarantees the security of information transmitted through the Internet. The PDA used to do the transmission, was connected to a microcomputer through the Microsoft Active Sync. Software.

In these areas, in case there were communication problems through the VPN, the PDAs data were copied to the computer and could, alternatively, be sent by CDROM, FTP or electronic mail to the central computer located in IBGE networks, in Rio de Janeiro.

The remaining collecting stations (about 4,400) did not have a personal computer; therefore, communication with the central base was done by a fi xed telephone line, through the use of a modem with Bluetooth technology, to communicate with the PDA. Initially, a connection between the PDA and the modem to be used in that station was established. In the beginning of the transmissions, the PDA software dialed to a fixed telephone line, which was bought from a national operating company with local taxes charged in the place of destination of the call, and, after the synchronization of the PDA and one of the computers in the census, Rio de Janeiro, the data were sent. In this case, the data were turned into cryptographic information in a private network and not through the Internet, as it was in the case of Computerized Collection Stations.

During the data collection period, the enumerators periodically transferred the data obtained to the central equipment in IBGE. At the end of the collection, it was sent to the corresponding supervisor, who: scanned the data collected with an audit software; finished the work in the census track; generated the return action to the field, or the conclusion of the collection in the census track; and sent the data to the central computer in IBGE, completing, this way, the collection process.

The transmission of data was done through remote connection close to the collecting area, which assured the integrity, security and quality of data transmission and its proper storage in the central computer of IBGE. Transmission controls were aimed at guarantee the data of the enumerator (who was responsible for the transmission) were kept with security in the central computer.

The partial or final data sent to the central computer of IBGE were always subject to a central plan of correction (data consistency or quantitative management plan), of which reports were transmitted to handheld computers (PDAs) of enumerators, in a way as to allow the individual control of fi eld operations by their own executors.
Therefore, besides scanning data with a local correction program, the PDA application software was also prepared to receive information about the quality of data transmitted and stored in the central computer.