



Hague, October 6th 2009



## **Spatial Information in Polish Official Statistics**

**Janusz Dygaszewicz**  
**Director of the Central Census Bureau**  
**Director of Programming and Coordination of Statistical Surveys Division**  
**Central Statistical Office**

### **INSPIRE**

Central Statistical Office (CSO) is actively participating in works regarding the implementation of the Directive 2007/2/EC of the European Parliament and of the Council of 14 March 2007 establishing an Infrastructure for Spatial Information in the European Community (INSPIRE).

In Poland the contact point (as mentioned in article 19 section 2 of the directive), responsible for the contacts with the European Commission in terms related to the inspire directive, is Surveyor General of Poland. In order to support works on the implementation of the directive, the Surveyor General of Poland appointed the INSPIRE Implementation Council. Tasks of the Council include: expressing opinions in matters of legislation, organization and in technical matters regarding the Polish Spatial Data Infrastructure (PSDI), which is created accordingly to the INSPIRE Directive. Furthermore, the Council recommends actions and projects of the inter-departmental matter in the field of geoinformation and supports Polish experts in committees and European Commission teams regarding INSPIRE.

As a result of actions of the national contact point and members of the Council, a draft of a spatial data infrastructure legislation act has been created. The act regulates basics of creation and operation of the Polish Spatial Data Infrastructure. PSDI will include all levels of public administration and serve all users of geographical information in Poland and in the Community.

Central Statistical Office was appointed as a leading unit in terms of two spatial data themes mentioned in appendix III: "Statistical units" and "Population distribution – demography". As a cooperating unit, CSO undertakes actions in terms of two themes from appendix I: "Administrative units" and "Addresses".

Along with activities in the matter of legislation, a testing team for appendix I themes' data specifications has been appointed. The testing process' goal is to analyze relations between data sets existing in Poland and themes specified by the INSPIRE directive, create a preliminary technical and economical analysis of the INSPIRE data themes implementation, design a schedule of works in the field of appendix I INSPIRE themes' data sets creation.

Unfortunately, the testing of appendix I themes' data sets led to a conclusion that there are no data sets in Poland that completely comply with the technical specification prepared by Thematic Work Groups. In most cases data is gathered by different, not cooperating public institutions, so none of these institutions is able to provide a complete set of objects and attributes. Steps are taken at the moment to achieve compliance of spatial data sets with the INSPIRE requirements.

Appointment of a Metadata Team, which is expected to prepare the technical specifications for creating metadata for appendix I spatial data sets and services within two years.

## **Address as a universal administrative registry data connector**

One of the main objectives of the official national registry of territorial division of the country - TERYT led by the President of the Central Statistical Office, is to provide unambiguous identification of objects with different levels of territorial detail, such as: voivodship, county, municipality, town, village, statistical district, census region, street, building and apartment. TERYT allows the collection of data for these spatial objects and provides the conditions for comparison and analysis, which is a very important factor in the implementation of the INSPIRE 2007/2/EC Directive of the European Parliament and the Council dated March 14, 2007 which establishes an Infrastructure for Spatial Information in the European Union (Journal of Laws, EU L 108, 25.04.2007, p. 1-13).

CSO is currently working on adding spatial information to the territorial identification registry. Administrative division of Poland already has its spatial representation in the form of the National Registry of Borders and Areas of the Country's Administrative Division. Steps are taken to create spatial data for the statistical division of Poland. Acquisition of such a data set is expected in the end of 2009. Address point acquisition for the whole country is currently in a pilot phase. So far 4 address databases have been built for 4 communes selected for the trial agricultural census, which takes place September 2009. Address point coordinates are acquired with help of following materials: basic (cadastral data and/or orthophotomap, address point sketches), referential (LPIS, Topographic Data Base, VMap Level 2).

The introduction of address points will allow a change of the existing system of spatial identification and will enable a transition from area allocation (census areas) to point allocation. This has crucial, revolutionary importance for geomatic usage in statistics. Changing the allocation will allow a more flexible aggregation of data collected in the national census, for any area.

It will also allow creation of microdata database with a spatial reference, which will enable geostatistical analysis of various effects such as:

- demographics (e.g. the average distance between children and parents within the country, voivodship, county, municipality, village, township, or street blocks or any other described area, the average distance from work, school, hospital, etc.)
- town and country planning (e.g. to assist in determining the borders of metropolitan regions, cities, development of spatial management plans)
- agriculture and environment (e.g. crop structure survey, environment contamination),
- the economy (e.g. study the effects of arduous road and industrial investments).

Allocating points with x, y coordinates will also allow independence from disruptive changes in the territorial division of the country, typically resulting in changes in census areas and the resulting arduous recounts. This will facilitate comparative analysis of time series irrespectively to changes in the division. An additional advantage is the aggregation of data in both the administrative division of the NUTS system as well as the GRID divisions developed within the EU GEOSTAT project.

## Census Realization Model

In the forthcoming censuses Poland, like most European countries, will implement a mixed model approach consisting of a combination of data from administrative sources with data acquired from statistical surveys. We believe this method to be the most effective and secure at the present development level of administrative sources and the degree of advancement of the methodological work enabling the use of administrative sources.

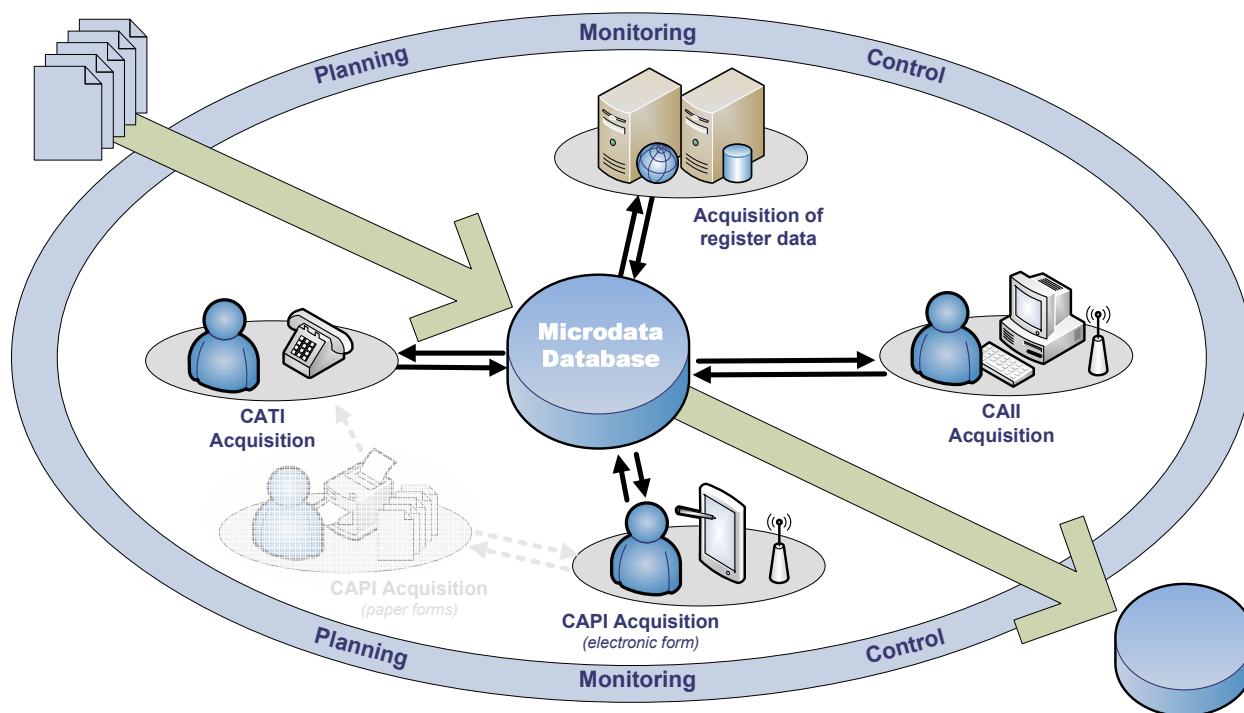
Poland, however - as one of the first countries in the world – will implement a fully innovative method involving the use of multiple modern census data collection techniques, **at the same time**. This especially applies to replacement of paper questionnaires with electronic questionnaires, which allow more efficient data collection. These include Internet technologies enabling self-enumeration over the Internet. The electronic questionnaire runs both online and off-line. A technology will be implemented in order to allow telephone interviews (CATI method). Also, census enumerators equipped with portable hand-held devices will use electronic questionnaires. Hand-held devices will enable census data collection, on-line transmission of data and the use of digital maps, which will eliminate the need for paper maps and situation sketches. The combination of digital maps and aerial photos with built-in GPS receivers is a revolutionary change in the development and census management before and during the census, it also enables multi-dimensional spatial analysis of the census results.

All of the above technologies will be used in the upcoming census, at the same time. This means, that the mixed method will use four channels for data acquisition:

- administrative registries
- Internet – self-enumeration (CAII)
- telephone interview (CATI)
- census enumerator visit with the use of hand-held devices (CAPI)

The last three channels will be supported by on-line electronic questionnaires.

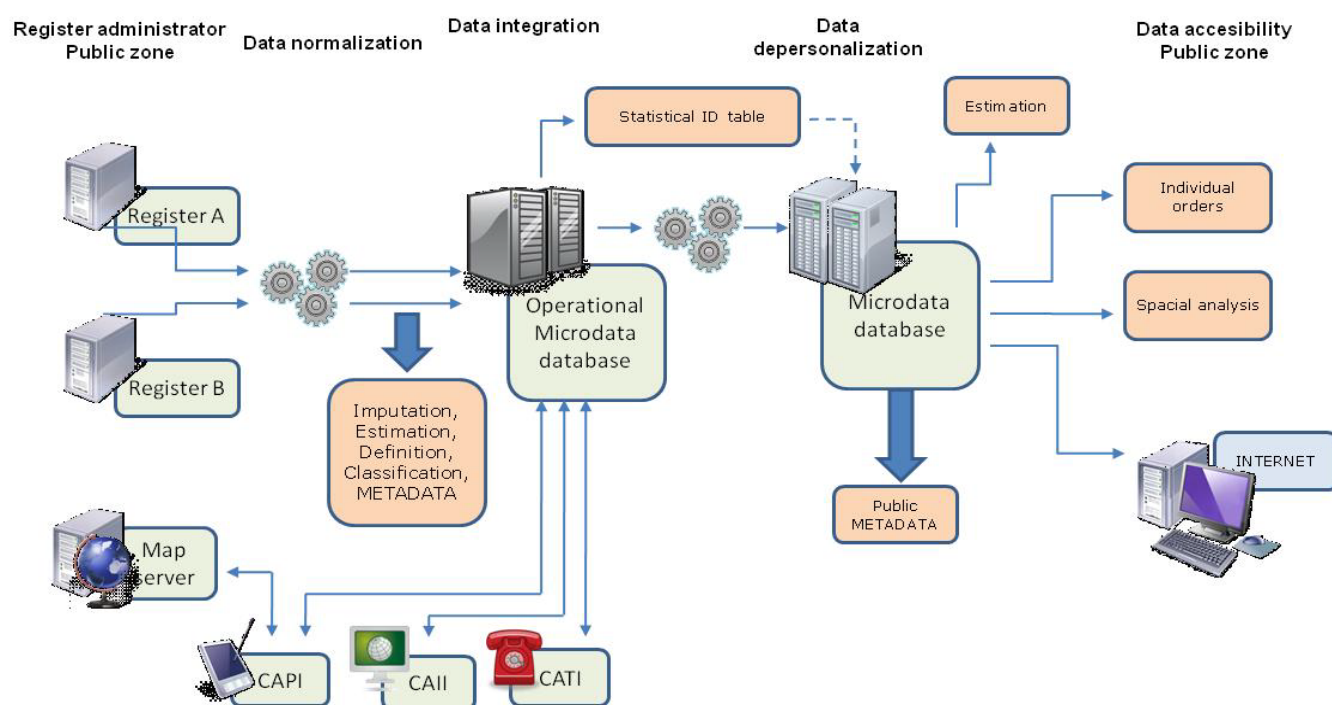
Primary data acquisition sources:



## Data processing

Data acquired from the four preliminary sources (described above) will be integrated in an Operational Microdata Database. This is one of the most important stages of census data processing, as it will decide upon the completeness of census data, its quality and the possibilities of data imputation and estimation. Data integration in the operational microdata database with regards to data from administrative registries is preceded by a very important stage of cleaning, standardization, de-duplication and normalization of data. Each information source is represented in the operating database, as a separate layer of data. The integration will rely on combining those layers basing on defined identifiers or key attributes. The Operational Microdata Database will allow advanced cleaning and repair of variables in the context of data available from all information sources. Ultimately, it will be possible to calculate, according to defined procedures and rules, the most probable values for individual census variables and transfer them to the target microdata database while depersonalizing individual data.

The microdata database will provide the basis for creating products such as census output tables, spatial analysis and basic result aggregates. It will also be a source database for the provision of individual data for scientific research, while maintaining the principles of statistical confidentiality. Microdata database, together with its accompanying metadata system will contribute to the public information system of the official statistics.



## Trial Agricultural Census

The census is preceded by a trial census. Its task is to comprehensively evaluate the technical, organizational and methodological solutions. Based on the experience from the trial census, we will be able to verify assumptions and modify those aspects of activities which may be necessary for censuses to be held in a smooth manner and with maximum efficiency.

A trial agricultural census is carried out at the moment. Three channels of data acquisition are in use by enumerators, telephone interviewers and supervisors.

Supervisors supported by the supervisor application, assign enumerators to census areas and assign specific farms for them to visit. Duties of a supervisor include: planning and monitoring the work of enumerators, census

completeness monitoring (verifying and accepting interviews, changing data acquisition channels when necessary), troubleshooting and handling emergency situations (real-time tracking of enumerators is supported by the application).

Enumerators are obliged to visit all farms planned to visit by the supervisor and interview the farmers using the electronic questionnaire on a mobile device. The devices are equipped with a map module, consisting of: ortophotomap, administrative and statistical division border and cadastral data. The module is used to visualize daily tasks on the map.

Interviewers' duties include: handling the CATI channel: telephone interviewing using the electronic questionnaire in the CATI application. Furthermore, interviewers deliver support to respondents using the self-enumeration method and schedule visits for enumerators.

Supervision of the census in the field is the task of Commune Census Bureau and Voivodship Census Bureau. Global supervision and support of the census is the task of the Central Census Bureau (central supervisors, IT experts).

## **Instead of a conclusion...**

Usage of new technologies, a new approach to censuses, the abandonment of paper, data collection through on-line electronic channels and the wide use of data from administrative registries requires enormous organizational, intellectual and financial efforts, which are applied into preparatory work. The support of the government administration, local government, as well as the society is also crucial. All involved in the organization of the census are aware of the importance of this task. Also, all along we believe in absolute success of censuses carried out using the latest technology.