

SEMINAIRE

E/CN.3/2000/10

SEMINAR

STATISTICAL COMMISSION AND
ECONOMIC COMMISSION FOR EUROPECONFERENCE OF EUROPEAN
STATISTICIANSDistr.
GENERALCES/SEM.43/18
14 February 2000

ENGLISH ONLY

Seminar on integrated statistical information
systems and related matters (ISIS 2000)
(Riga, Latvia, 29-31 May 2000)

Topic II: Planning and management of statistical projects

CENSUSES 2000: MAIN STRATEGIES AND IT SUPPORTING PHASES IN SLOVENIA**Contributed paper**Submitted by the Statistical Office of the Republic of Slovenia¹**I. INTRODUCTION**

1. The censuses in Slovenia have been supported by up-to-date Information Technology (IT) during the last few decades. Each time when the census turned around a decade, new possibilities of IT were introduced. Over time, the use of IT has become one of the key elements in the efficient organisation of the whole census campaign. The role of IT has increased to cover almost all phases of work from Census preparation to data dissemination. The main challenges nowadays lie in:

- the integration of data from different administrative and statistical sources;
- the reduction of the burden for respondents;
- the integration of operational processes;
- dissemination of the results to a larger public than ever before.

II. ORGANISATION OF THE POPULATION CENSUS AND THE CENSUS OF AGRICULTURAL HOLDINGS

2. Censuses in Slovenia are carried out regularly. The Statistical Office of Slovenia has extensive experience, especially with population censuses, since we

¹ Prepared by Milan Katic.

have been implementing them independently for decades.

3. In the near future, two major censuses will be organised: Census of Population, Households and Housing (planned for 2002) and the 2000 Census of Agricultural Holdings. Both these censuses will, to the greatest possible degree, use the existing data stored in various administrative sources. We have at our disposal high quality information infrastructure consisting of various administrative registers and databases, such as Population Register, Business Register, Register of Territorial Units, Tax Register, Employment Register, etc.

4. These registers contain unique identifiers that allow to link data with other registers. This enables us to use their contents for different purposes, which is one of the strategic aims of the Statistical Office. Linking Census data with administrative registers was done to a certain extent already in 1971. During the last census in 1991, some census data were actually obtained from administrative sources.

5. Good results of this approach in previous censuses, as well as further development of administrative registers were the basis for the *strategic decision* that future censuses will be to a great extent register-based.

6. This strategy has implications for the role and function of IT. It is no longer focused merely on technical operations, such as data transfer from paper questionnaires to electronic media, editing, tabulation, etc., but on multipurpose use and integration of different data sources.

7. However, the successful implementation of this task requires thorough preparation, as well as good management and organisation of work. As a rule, special project teams are set up for each individual type of census. These project teams are composed of experts from subject matter departments and IT experts. The project team is directly responsible and reports regularly to the Director General of the Statistical Office.

8. The key phases of implementing the 2000 censuses in which the role of IT is especially important are briefly described below. The main IT supporting functions in individual phases of the Census are listed in the following table.

Main Census phases	Main IT support functions
Conceptual phase	<ul style="list-style-type: none"> - general functionalities of IT; - possible data integration from different administrative sources; - possible functional integration of the census phases and tasks; - presentation and dissemination of the results.
Census preparation	<ul style="list-style-type: none"> - collection of personal and other relevant data from the existing sources; - preparation of the supporting documentation (charts and maps of census districts, lists of houses and persons);

	- pre-printing personal data on the questionnaires.
Data entry, editing and archiving	- functional integration of individual phases; - scanning and archiving of the questionnaires; - OCR (Optical Character Recognition); - data editing with direct access to the scanned questionnaires; - design and implementation of the microdata base; - archiving of the microdata.
Data aggregation and analysis	- design and implementation of the census macrodata base (including aggregation on different levels); - integration of aggregates from other sources; - support to analytical functions (especially OLAP);
Output and dissemination	- tabulation; - publishing the results (including in electronic format); - Internet and Intranet dissemination.

III. CENSUS PREPARATION PHASE

9. Extensive and thorough preparations are needed. The various organisational and technological tasks can only begin after the problems regarding contents and methodology are solved, and the general deadlines for the implementation are set (a special legal act is adopted).

10. One especially important task is the specification of relevant administrative sources and data to be collected from these sources. These are:

- personal data that will be pre-printed on census sheets (in the implementation phase these data will be checked by the interviewers and, if necessary, corrected);
- data from the Register of Territorial Units, on the basis of which detailed maps of census districts and the list of houses in the census districts will be prepared (using geo-referential model of Slovenia).

11. A complete list of persons living in these houses will be prepared from the Population Register. In addition to basic census questionnaires, this is the main documentation that interviewers will have at their disposal. The quality of these documents should ensure the high quality of interviewing and various subsequent census controls.

12. The IT support for setting up the pre-census databases in which personal data and data on census districts will be integrated, and the preparation of the

above-mentioned documentation will be partly outsourced.

IV. DATA ENTRY, EDITING AND ARCHIVING

13. The general strategy of data entry includes scanning census questionnaires and OCR technology of data capture. Scanned documents (images) will be stored in the *documentation database* and archived. The digitized records obtained using the OCR technology will be stored in the *input database*. In later phases of data editing, the interactively accessible database of scanned documents will be used as the basis for correcting the data which will further speed up data editing.

14. From the point of view of resources, this is one of the most demanding tasks of every census in which most data (despite the use of existing sources) are collected in the classical way. It requires very detailed planning and specification of individual sub-phases and even of individual activities. We will evaluate the possibilities to outsource the scanning and data capture using OCR for the population census, since we still have some time before taking the decision.

15. In these phases, IT plays a universal role. *Integration of scanning and OCR processes* is especially important as well as interactive use of the documentation database in the editing phase. The data editing (control and correction) will be performed by the Statistical Office using our own applications. During pilot censuses, the Blaise system for editing was successfully tested. Blaise will also be used for the Census of Agricultural Holdings in June 2000.

V. AGGREGATION AND ANALYSES

16. Aggregation will be performed on the basis of clean data in the microdata base and the macrodata base will be set up. The level of aggregation is the statistical district as the territorial unit that enables aggregation into larger territorial units such as settlements, municipalities, regions and the entire country.

17. The macro database of aggregates will contain, in addition to census aggregates, other aggregate data from administrative and statistical sources (e.g. 1998 and 1991 Censuses). Thus, a single database will be set up for various analytical purposes and for post-census research of socio-economic phenomena.

18. In this phase the IT fulfils an *integrative* role:

- the integration of census aggregates with data from other sources;
- supporting on-line analytical functions with corresponding OLAP tools which have developed extremely rapidly in the last few years.

VI. OUTPUT AND DISSEMINATION

19. It is planned to implement a tabulation system, which will produce all

census output tables, as well as publication and dissemination of census results both in paper form and through electronic media. It is also planned to prepare an Atlas of Slovenia where the main census data will be presented on maps and charts. The highest level of aggregate data will be accessible via Internet.

20. In this phase, the role of IT has changed most compared to previous censuses. The census results will be made available via Internet and disseminated in electronic form in addition to printed publications. This opens up possibilities for a large-scale use of data for analyses and processing.

VII. CONCLUSION

21. Implementation of censuses is a regular activity in Slovenia. The Statistical Office of Slovenia has extensive experience, especially in carrying out population censuses, since for decades we have been implementing them on our own. Good quality information infrastructure, such as various administrative registers and databases is available for successful and efficient implementation of censuses. The above-mentioned registers contain unique identifiers that can be linked with other registers, which allows to use their contents for different purposes. This is one of the strategic aims of the Statistical Office. The current level of the information and technical infrastructure allows to efficiently implement such approaches.