

**STATISTICAL COMMISSION and
UN ECONOMIC COMMISSION FOR
EUROPE (UNECE)**

**STATISTICAL OFFICE OF THE
EUROPEAN COMMUNITIES
(EUROSTAT)**

**CONFERENCE OF EUROPEAN
STATISTICIANS**

Joint UNECE-Eurostat Work Session on Population Censuses
Organized in cooperation with UNFPA
(Geneva, 23-25 November 2004)

Session 3– Invited paper

**TECHNOLOGIES USED BY ECE COUNTRIES IN THEIR 2000 ROUND OF
CENSUSES**

Submitted by Statistical Division, UNECE*

BACKGROUND AND SUMMARY

1. The fast developing of emerging technology is changing the way censuses are carried out. The 2000 of census round witnessed the introduction of many new technologies in the census, from the scanning replacing manual data editing to the elaboration of digital maps and the use of GIS. For some countries some of this technology was still used on an experimental basis (as the use of scanning) for others these were used as consolidated practices. This note reports the practices used in countries for the monitoring of the census operations, mapping and data processing and editing in the 2000 census round. This analysis is based on a questionnaire that ECE sent to all its member countries on the practices of the 2000 census¹. The note also makes some suggestions on the issues that should be covered in the ECE Recommendations for 2010 round of censuses on the use of technology (as suggested by the Steering Group on Population and Housing Census). The aim is not to recommend a specific software or a technology but rather to provide guidelines on how countries can best choose what methodology to use.

* Paper prepared by Angela Me and Paolo Valente from the Social and Demographic Statistics Section. The views expressed in this paper are those of the authors and do not necessarily reflect the views of the United Nations.

¹ Among the 56 countries (11 ECE members plus Australia) that received the questionnaire, 44 countries responded.

1. MANAGEMENT OF THE QUALITY OF CENSUS OPERATIONS

2. Among the 44 countries that replied to the ECE questionnaire, 27 reported the use of a software to support the management of the census². 10 of these countries developed an ad-hoc software while 12 reported the use of commercial software. Among the commercial software used, the most common is Microsoft project (11 countries). Table 1 presents the list of software used as project management tool. 18 countries reported that did not use the computer to monitor the quality of the census operations.

Table 1. Type of software used by countries as census management tool.

Armenia:	CSPro and IMPS
Australia:	Microsoft project
Belgium	Formiris 2.7
Canada	Suretrak, Primavera, MS Project
Croatia	Microsoft Project
France	PMW one by process
Georgia	Microsoft Project
Greece	Oracle, SQL and self-developed
Italy	Ms Project
Kyrgyzstan	Client-Server
Latvia	Ms Project
Lithuania	Ms Project
Norway	Microsoft Project
Portugal	Microsoft Project and and self-developed system (SIGINE)
Russian Federation	Self-developed firmware "Perepis'
Spain	DIA for automatic debugging and imputation
UK - England and Wales	Ms Project

2. DATA ENTRY

3. As it is shown in Table 2, about half of the countries (23) that responded to the ECE questionnaire reported the use of manual data entry. About 25 countries used OMR (Optical Mark Recognition) and 14 countries used OCR/ICR (Optical Character Recognition/Intelligent Character Recognition). 8 countries³ reported the use of internet for data entry. Some countries used more then one technology depending on the organization of the census particularly if the data entry was a decentralized operation. For some of the countries that did not use manual data editing, it was the first time that they used technology such as scanning or internet. For some of the countries the use of new technology was an improvement in the timeliness and quality of the data, but for some countries the use of new technology was a challenge. What is the best technology to use depends on the national circumstances. Issues such as labour cost and capacity of the NSO to manage the technology in terms of human and technical resources should be

² The 17 countries that did not report the use of a project management software are: Austria, Azerbaijan Republic, Bulgaria, Cyprus, Estonia, Ireland, Israel, Luxembourg, Netherlands, Romania, Slovakia, Slovenia, Switzerland, The FYR of Macedonia, Turkey, Ukraine

³ Belgium, Canada, Finland, Italy, Norway, Spain, Switzerland, United States. It is not clear from the answers given to the questionnaire if internet was used to collect information from the respondents or to facilitate the operation of data entry in the national statistical office.

considered before taking decisions on what technology to use. The latest technology is not always the best solution. Actually for NSOs where the resources (particularly human resources) are limited it may be better to use solutions that have been widely used by other countries and where the procedures and the actions to take when problems arise are clearer and well tested. Looking for example at the duration of data entry operations it can be noted that the most advanced methods such as OCR and OMR/ICR do not always assure a faster operation in respect to manual data entry.

Table 2. Number of countries that used different data-entry methods⁴

Data-entry Method	Number of countries
Keyboard	23
OMR	14
OCR/ICR	25
Internet	8
Other ⁵	3

4. Regardless of the method used to entry the data, there is always a margin of error. Double entry operations help to monitor the errors occurring during the entry operations and it allow not only to fix possible consistent errors but also to give a measure of the quality of the data entry. Less then half of the countries reported to have used double entry operations. Table 3 shows the countries that used double -entry operations according to the method used to entry the data and the percentage of data entered twice

Table 3. Countries that used double -entry operations with the methods used for data-entry and percentage of data double entered.

Country	% data double -entered	Data-entry Method
Armenia	35	Keyboard
Azerbaijan Republic	10	Keyboard
Bulgaria	20	Keyboard
Czech Republic	2	OMR
Canada	20.54	Keyboard
Italy	5	OCR/ICR
Kyrgyzstan	10	Keyboard
Malta	100	Keyboard
Poland	5	OCR/ICR
Slovakia	25	OCR/ICR
Ukraine	1	OMR
UK - England and Wales	100	OMR
UK - Northern Ireland	100	OMR
United States	5	OMR

⁴ The same country may be counted in more then one method since countries reported the use of more then one method of data entry.

⁵ This includes the comparison between values optically recognized and values in the population registers (used in Israel).

5. More than double entry operations countries used real-time error control. 40 countries reported to have run such a control.

Table 4. Number of countries according to the data entry method used and the duration of the data-entry operation⁶

Data entry method	Duration of data entry (in months)	Number of countries
<i>Keyboard</i>	0-4	0
	5-6	8
	7-9	2
	>9	1
<i>OCR</i>	0-4	1
	5-6	1
	7-9	0
	>9	0
<i>OMR/ICR</i>	0-4	3
	5-6	4
	7-9	4
	>9	1

6. Looking at the software used for data entry it can be noted that more than half of the countries (27) developed their own-application using the language reported in Table 5. Less countries (18) used commercial products (listed in Table 6). Two countries used free-non commercial-products: Armenia (CSPю) and Hungary (Bull-LaPoste, French product). Some countries used a mix of own-developed applications and commercial applications.

⁶ Only countries that indicated one method of data entry were included in the table

Table 5. Language used by countries that developed their own-application for data entry

Country	Language used to develop own-application for data entry
Armenia	Visual Basic
Albania	Visual Basic
Australia	Microsoft C
Azerbaijan Republic	Visual FoxPro
Belarus	Visual FoxPro
Belgium	Formiris and Informix
Bulgaria	Visual Basic
Canada	Census Automated Control System developed by Canada Revenue Agency
Cyprus	AFPSPRO developed by private company
Georgia	Power Builder
Greece	Visual Basic, SQL, PL-SQL
Italy	ORACLE FORMS
Kyrgyzstan	Client-Server architecture, Query Language - SQL and tools for the development of DELPHI software, Pentium Pro Server
Malta	FOXPRO
Norway	SAS
Portugal	C++
Romania	MS visual FoxPro
Russian Federation	C, C++, Visual Basic Script
Serbia and Montenegro	Visual Basic, Windows, NT, SQL, PL-II and Access The FYR of Macedonia: VISUAL BASIC

Table 6. Commercial products used by countries for data capture

Country	Commercial product
Australia	IBM Intelligent Form Processing Eyes&Hands
Austria	IBM Intelligent Form Processing; STAR recognition software (OCE)
Belarus	DBMS Oracle 8
Croatia	IBM Intelligent Form Processing
Estonia	ReadSoft AB ICR/OCR software
Georgia	Readsoft Eyes&Hands for Forms
Ireland	Bespoke System build on AFP SPRO (Top Image Systems Israel)
Latvia	Eyes and Hands
Lithuania	monsun/2
Portugal	Floware form Plexus
Slovakia	AFPS pro
Slovenia	READSOFT
Switzerland	Kodak capture
Turkey	AFPS -PRO
Ukraine	E&h Scan Sweden
UK - England and Wales	TMS Sequoia Formfix (OMR)+CGK Recostar (OCR)
UK - Northern Ireland	TMS Sequoia Formfix (OMR)+CGK Recostar (OCR)
United States	Lockheed Martin (DCS2000)

3. CODING

7. The countries in the region were still heavily relying on manual coding: about half of the countries (21) reported the use of code books to do manual coding. All these countries were from East of Europe or CIS. The number of countries increases to 23 for computed assisted coding and to 28 for automatic coding. As for data entry, countries often use a combination of the three methods.

8. The market of commercial packet for editing seems less advanced then the one for data entry. Only 6 countries reported the use of ready-packages for data coding (see Table 7) while 26 countries reported the development of their own package.

Table 7. Packages used by countries for computer-assisted or automatic coding.

Country	Package used for data coding
Croatia	ACTR - Statistics Canada
Czech Republic	IRIS sw
Ireland	Precision data coder for Occupation coding
Italy	ACTR BY STATISTICS Canada
UK - England and Wales	ACTR (textual responses), MATCHCODE (address coding)
UK - Northern Ireland	ACTR (textual responses), MATCHCODE (address coding)

4. EDITING

9. All the countries that replied to the ECE questionnaire reported that they performed computer-supported editing. Among these 35 countries included imputation. Countries that did not perform any imputation are small countries or countries mainly form CIS and East of Europe. Among these 35 countries, only 23 generated statistics on imputation rate by variable.

5. MICRODATA DATABASE

10. Almost countries set up a database with census microdata. Table 8 reports that type of database developed.

Table 8. Type of database developed for microdata

Type of database	Number of countries
High level (Oracle, SQL Server, ...)	34
Desktop	5
Statistical (SPSS system file, etc.)	13
Demographic (REDATAM, etc.)	2
Other	5

6. MAPPING

11. 39 out of 44 countries reported the use of digital maps for the 2000 census round while 40 countries reported the use of GIS. Unfortunately from the data available it is not possible to understand at what stage of the census and what for the GIS technology was used.

Suggestions for a table of contents of a new session of the ECE Recommendations on technology

1. Data entry (reference to the existing UN handbooks)
 - 1.1. Review of options
 - 1.1.1 Brief Description
 - 1.1.2 Advantages and Shortcoming
 - 1.2 Conditions that need to be in place in order for each solution to be successfully applied (What make a successful solution?)
 - 1.3 The need to monitor data entry errors and measure the margin of error: double-entry operations
2. Mapping
 - 2.1. The use of GIS and GPS in improving coverage
3. Data coding
 - 3.1. Key issues to consider and best practices
4. Data editing (reference to the UN Handbook on data editing)
 - 4.1. Key issues to consider and best practices
5. Sub-contracting (reference to the UN Handbook on Census Management)
6. Methods to assure confidentiality
 - 6.1 Data dissemination
 - 6.2 Microdata
