**UNITED NATIONS STATISTICAL COMMISSION and
ECONOMIC COMMISSION FOR EUROPE (ECE)
CONFERENCE OF EUROPEAN STATISTICIANS**

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

**ORGANISATION FOR ECONOMIC
COOPERATION AND DEVELOPMENT (OECD)
STATISTICS DIRECTORATE**

Joint ECE/Eurostat/OECD Meeting on the Management of Statistical Information Systems (MSIS)
(Geneva, 17-19 May 2004)

Topic (i): Web technology in statistical information systems

WEB-SUPPORTED STATISTICAL DISSEMINATION PROCESS SERVING THE USERS

Supporting paper

Submitted by Statistical Office of the Republic of Slovenia [1]

Summary

I.   INTRODUCTION

1.   The paper will present two aspects of using web technology in statistical information systems: web technology in the statistical process and web technology in dissemination of statistics. Both uses will be presented on practical cases, developed in the Statistical Office of the Republic of Slovenia (SORS).

2.   The corporate (statistical) data warehouse is used in SORS as a target common data management concept. This concept is very closely connected with the presentation of the statistical data and metadata on the web. We have developed some web supported solutions both for the statistical process and for dissemination.

II.   COMMON ANALYTICAL ARCHITECTURE IN SORS

A.   Main elements of the architecture

3.   The statistical data warehouse is closely related to metadata systems and dissemination systems. The classification database KLASJE contains all classifications used in SORS and provides automatic extraction of classification related metadata (categories, levels, hierarchies, etc.) which are used in codelist tables, star scheme dimension tables and cube dimensions of the statistical data warehouse. The classification database was built on the basis of New Zealand classification model (CARS) and has been in production since 2000.

---

[1] Prepared by Matjaz Jug (matjaz.jug@gov.si).

GE.04-30300

4.   The meta database METIS contains structured metadata about surveys, questionnaires, questions, observation objects, variables, their valuesets and values, publications, activities, etc. The database is finished and is in the process of deployment. Some parts are already in test production.

5.   A new dissemination database, based on Swedish pseudo-OLAP file format PC-AXIS, was built in STAT2000 project and is the primary consumer of the data from the statistical data warehouse.

III.   WEB TECHNOLOGY IN THE STATISTICAL PROCESS

A.   Statistical process from tabulation to publication on the web

6.   Data extracted from the statistical data warehouse and other sources are aggregated into final multidimensional matrixes, stored in PC-AXIS format. The PC-AXIS file format includes many important types of metadata, for example administrative metadata (refresh date, matrix name, etc.), methodological metadata (footnotes, indicators, etc.). The preparation of the dissemination matrix should be supported with a powerful user interface, so we are using the stand-alone application PX-Edit with broad functionalities for importing, manipulating and maintaining the multidimensional matrixes.

B.   Web form as a tool for submitting final statistical data

7.   Finished matrixes have to be submitted by the responsible subject-matter staff on the central dissemination server into the corresponding folder. The web form used for this purpose offers some very useful possibilities. Firstly, submission of the matrix is easy and fool-proof because of predetermined fields, checked by the metadata already in the database. Some administrative metadata, which are not included in the PC-AXIS format (for example the identification number of the related printed publication) could be entered through the form. Files with multidimensional matrixes and related metadata are placed in the corresponding directory where some important metadata are extracted from the PC-AXIS file and placed in the relational database together with metadata, which were entered directly on the form.

IV.   WEB TECHNOLOGY IN DISSEMINATION

A.   Presentation of different types of information on the web

8.   SORS's Internet site (http://www.stat.si) offers different levels of presentation of statistical data. Main figures (actual indicators) are presented in a very simple but powerful user interface with a short methodological comment, possibility to show the figures in the chart and do the recalculation of indices between different time periods. Printed publications are presented mostly using PDF and static tables easy to download and print. More detailed data are presented as multidimensional matrixes with the possibilities to filter the data according to variables and their values, pivot the table structure and download the resulted table in a different file format. Some advanced charting possibilities are possible and the presentation of the data in maps, there are direct links to the document with the methodological explanations and some additional formalized metadata.

B.   Putting all together: linked dissemination services

9.   Tightly integrated web sites with smart links to really related topics and contents are state of the art today. Here is the point where different phases from the statistical process find their expression: metadata collected using stand-alone applications and web forms serve as a source for automatically managed links on the interconnected web site. Links should be exactly where the users need them to effectively navigate between different presentation types of data and different types of metadata.

10.   What are the future plans? Web services are a very good basis for building a well interconnected data service both internally and over the Internet.

- - - - -