A GOOD PRACTICE OF COOPERATION BETWEEN SOCIAL SCIENCE DATA ARCHIVES AND A NATIONAL STATISTICS OFFICE

The Slovenian example

assist2013
May 28 – May 31 in Cologne | Germany
The cooperation as an initiative of the DwB project

- has been going on since the establishment of the Slovenian Social Science Data Archives (ADP), started with distributing anonymised SORS microdata and structured metadata - surveys such as Labour Force Survey, Census 2002 etc.

- less intense cooperation between 2002 and 2011, no newly distributed SORS microdata; increased cooperation since the start of the DwB project

- the main goal of the DwB project is to support equal and easy access to official statistics microdata for the European Research Area, our cooperation concentrated on the research environment in Slovenia

- the products of our work could be used in modified form to achieve DwB project goals easier, especially working on Work package 5 tasks

*The Slovenian example*
The purpose and the aim of our cooperation

- one of Social Science Data Archives’ missions is to provide quality microdata and metadata to different groups of data users: under-graduate students, postgraduate students, young researchers at faculties, research organizations, registered researchers – we cannot concentrate on the serious researchers’ needs by mostly distributing Public Use Files

- ADP can also help SORS with its experience and expertise in the field - researchers can benefit in the long term

- SORS could use some additional human resources for microdata access service

- ADP has additional knowledge of DDI metadata standard, students’ needs, experience with training of researchers

- both organizations benefit from established additional communication channels for microdata usage promotion

The Slovenian example
Forms of cooperation – what we do

• **detailed microdata**: intensive work in SPSS on deindividuated microdata - preparing easy-to-use microdata in the safe center environment, adding variable and value labels, missing values to datasets, additional logical control, deleting unneeded variables, connecting variables in databases to codebooks

• **PUF microdata**: preparing Public Use Files, later distributed by ADP – some of our target groups are not familiar with SORS microdata, the majority of other researchers would also benefit from a simpler access to moderately anonymised microdata

• metadata: analysis of possible automatic harvesting of metadata from SORS metadata systems, preparing structured metadata for researchers, counseling on inclusion of DDI standard to SORS procedures

• activities to **promote official statistics microdata scientific use** – conferences, mailing lists, websites, social network sites

*The Slovenian example*
Preparing deindividualized microdata – how we do it

Forms of cooperation

The Slovenian example
Preparing Public Use Files – how we do it

- in cooperation with SORS Sector for General Methodology and Standards
- developing new procedures, evaluating quality and applicability of anonymisation software and its technics
- anonymisation procedure starts with initial anonymisation which follows Eurostat LFS anonymisation criteria (in SPSS)
- followed by calculating individual & global risk ($R! - sdcMicro$)
- calculating strata allocation, based on individual risk averages by strata ($R! - bethel$)
- stratified sampling, based on the inclusion probability of a certain case ($R! - sampling - samplecube$)
- sample weights recalculation is done at the end

The Slovenian example
METADATA – how we do it

• the analysis of SORS metadata systems and other possible metadata sources was made, we realized that some bits and pieces of metadata could be harvested with involvement of several employees from some databases, applications, e-versions of documentation (quality reports, methodological explanations, databases, questionnaires etc.), **no automatic harvesting is possible**, at least for now

• SORS’s been working on **improvements of its microdata and metadata documentation systems**; as Social Science Data Archives have extensive experience with DDI standard, counseling SORS work groups on metadata standards

• addressing researchers’ needs, structured metadata are being prepared in the form of **study descriptions**, DDI scheme is used – methodological, file description, data description, publication, other material etc. metadata fields are included

*The Slovenian example*
## Labour force survey, 2010: Annual data - no confidentiality restrictions

<table>
<thead>
<tr>
<th>Study description</th>
<th>Data description</th>
<th>Materials and publications</th>
<th>Download data</th>
</tr>
</thead>
</table>

**ADP - IDNo: ADS10**

**Main author(s):**
Statistical Office of the Republic of Slovenia

**Data file producers:**
SORS - Statistični urad Republike Slovenije = Statistical Office of the Republic of Slovenia (Ljubljana, Slovenia; 2012)

**Funding agency:**

1. Based on the Law on National Statistics, all the activities, listed in the Annual Program of Statistical Surveys 2010, are financed by the National Budget of the Republic of Slovenia.
2. European Commission with funds to be used for methodological improvements (occasionally).

**Series:**

- ADS/Labour force survey

Labour Force Survey is the biggest official statistics household survey in the Republic of Slovenia. A purpose of the survey is to collect information about current state and changes in employment and unemployment and economic inactivity in Slovene labour market. It provides data on the size, structure and characteristics of the active (employment and unemployed persons) and non-active population of Slovenia. Until 1996 the project was named Workforce Survey. As the development project was launched at the Institute of Social Sciences at the Faculty of Social Sciences. The development phase lasted until the end of 1999. After 1993 it is part of Official National statistical research program. From that time on the data has an official status. From 1998 on Labour Force Survey is conducted under auspices of the Statistical office of the Republic of Slovenia. In years 1993 and 1994 the survey was held together by Employment office of Republic and the Statistical office of the Republic of Slovenia. Labour Force Survey is conducted in accordance with the instructions of International Labour Organization (ILO), adopted by the 13th Conference of Labour Statisticians; and in accordance with the requirements of the European Union Statistical Office (Eurostat), based on harmonized survey of the European Union. Comparability of data with other countries and across time is thus

**Complete study description**

**Access to data:**

- Study description | Download data

**Status of the study:**

- 4: Full Study description and XML DDI Codebook. Data description with full questionnaire text.

**How to quote this study?**


**Related studies:**

- ASK84 - Workforce survey 1994
- ASK91 - Workforce survey 1991
- ASK92 - Workforce survey 1992
- ASK93 - Workforce survey 1993
- ASK94 - Workforce survey 1994
Availability of prepared metadata – Nesstar

The Slovenian example
Availability of prepared metadata – on-site laboratory, remote access

The Slovenian example
Availability of prepared metadata – html browsing document

The Slovenian example
Availability of prepared metadata – html browsing document

The overview tool in front of you is prepared to be used for browsing metadata for the listed researches of the Statistical Office of the Republic of Slovenia. To access microdata you need the approval from the Data Protection Committee. Then you can use databases and conduct statistical analysis. The tool is prepared both in Slovene and English and the language can be changed by clicking on one of the desired languages on the button "SLO" | "ENG".

Surveys are sorted by series. On the left side you can find a list of surveys with prepared metadata. You can select what metadata you would like to search by clicking on one of the buttons on the right side. The following information could be selected:

Codebooks used by year: The list of codebooks, available to researchers to clarify the values of individual variables, is presented by the year of conducting a Labour Force Survey. Codebooks used have changed between different periods of conducting a survey.

Variable comparison by year: The list of all variables of the Labour Force Survey is presented. Two separate lists are presented, from 2001 to 2004 and from 2005 to 2011, because of the major changes between the surveys LFS 2004 and LFS 2005. The document is meant to be used as a basis for building time series.

Study description: The whole study description is available; DDI 2.0 metadata standard is used. The description includes a big spectrum of metadata, from the information on the author of a research to the information on collecting data and research documentation, required for a researcher to be able to use a database appropriately.

Data description: Data on the variable level, frequencies or descriptive statistics, are presented. Variable values with low frequencies are hidden to protect privacy of a research respondent.

Other documentation: The documentation, required for a researcher to be able to use databases efficiently, is presented. It includes questionnaires, explanatory notes, codebooks and important scientific publications as a result of a statistical analysis of the research data. The mentioned documentation could be easily accessed by clicking on the hyperlinks in the documents.

The Slovenian example
Results of the cooperation and challenges of the future

- **Labour Force Survey 2001-2011** detailed deindividualized microdata, structured metadata, survey documentation gathered, available in the on-site laboratory or by the remote access, register **data** on public expenditures to support pensioners are being prepared; preparation of other microdata and metadata will follow.

- Labour Force Survey 2010 **Public Use File is being prepared**; PUF files will be prepared parallelly to work on detailed microdata.

- Work in different **metadata** fields has produced some initiatives and development of the systems; further development is planned.

- Becoming familiar with organizational policies, procedures, work specifics, working with different SORS departments, developing procedures, **low input-high output model** of cooperation – all that will make work significantly easier in the future.

- Listening to researchers and helping them working with microdata.

*The Slovenian example*