

Access to official statistical micro data at the Statistical Office of the Republic of Slovenia and cooperation with the Slovenian Social Science Data Archives

Irena Krizman

Statistical Office of the Republic of Slovenia, irena.krizman@gov.si

Enabling researchers to reuse statistical micro data is one of the most important elements of the Statistical Office of the Republic Slovenia (SURs) strategy where we stated our will to provide a variety of products and services to different user groups. The poster presents the cooperation SURs has established with the Slovenian Social Science Data Archives (ADP) and researchers, which has been intensified in the last two years as a consequence of being part of the European Commission Framework 7 Programme »Data without Boundaries (DwB)«. Different elements of the access process at SURs are presented, such as data availability, application, internal procedures and modes of transmission, as well as cooperation with the ADP, such as shared preparation of public use files, preparation of metadata in compliance with the Data Documentation Initiative (DDI) standard and promotional activities. Cooperation provides better products and services for researchers, builds new expertise at SURs and the ADP, and increases the use of statistical micro data. Our practice is seen by many statisticians, researchers and archivists as a good practice. The DwB project provides a unique opportunity to pass our experience to the international statistical and research community.

Key words: access to micro data, metadata, DDI, cooperation, DwB

1. Background

As other national statistical offices (NSOs) SURs is also facing a growing demand for micro data to be accessed within the country as well as transnationally. Regardless of the variety of tabulated data we provide through our dissemination system, there are analyses that can be done only with micro data: analyses of relationship, hierarchy, geographically specific analyses and longitudinal analyses.

There have been important developments in a number of NSOs in the last decade. In general, more and more NSOs provide access to confidential micro data for research purposes. The procedures are less demanding than before but are still considered very demanding and not user friendly by researchers.

NSOs put the resources to organise the on-site laboratories and remote access facilities. The refusal of access to micro data is many times advocated by legal barriers - even though the challenge is more in the practices applied than in the laws which usually contain general provisions or are silent on the issue. Another important challenge is shrinking resources. It encourages NSOs to find the partners outside official statistics to help to accomplish the research community expectations. The DwB project gives a stimulus for better cooperation with other institutions such as data archives. The cooperation has shown promising results in some countries, including Slovenia. It is also important to mention that there are several international initiatives dealing with micro data access within the European Statistical System (ESS), the OECD, the UN

Statistical Commission and the UNECE. A globalised world needs open borders also in the field of research. In June 2013 a full day CES Seminar on Micro Data will be organised by SURS in cooperation with USA and New Zealand Statistical Offices. (CES Seminar on Micro Data 2013).

The paper presents the good practice regarding micro data access organised by SURS with the focus on cooperation between SURS and the ADP. With the increased number of requests for national and trans-border access and limited resources available, SURS and the ADP recognised the need to deepen their cooperation. One of the important features was also to enter jointly the DwB project. The paper also presents the main drivers, objectives and results expected through the experience of SURS and the ADP active participation in the DwB project.

2. Micro data access at SURS

SURS has nearly a 30-year-long history in providing micro data access. But due to huge demand for census of population and housing micro data, from 2004 on we have upgraded the system which was before neither transparent nor properly managed to a system which is transparent for data providers and researchers (http://www.stat.si/eng/drz_stat_mikro.asp). Today researchers enjoy a wide variety of datasets available, they can access data by different means and internal procedures are well managed, standardised and transparent to employees.

In the EU peer review process SURS's practice regarding micro data access was evaluated as one of the good practices in the European Statistical System. The report published on Eurostat website http://epp.eurostat.ec.europa.eu/cache/ITY_PUBLIC/PEER_REVIEW_SI in 2007 stated "Access to micro data for research - SURS provides carefully controlled access to micro data, according to well-defined Protocols. The research community accepts the obligations placed upon it. An on-site laboratory with a stand-alone computer is supported by SURS staff, who ensure that no disclosive information is removed from the Office."

Researchers enjoy a wide variety of datasets available ranging from population, housing and agriculture censuses, social surveys, business surveys and they can use and link all administrative records we use for the production of official statistics.

In the on-site laboratory primary protected (standard identifiers are removed) micro data are available. Researchers can link and analyse them, while final output is checked against confidentiality by SURS. Usage of the on-site laboratory is free of charge. After 2007 researchers have been offered a possibility to work from their locations and use the remote access facility so they don't need to visit the on-site laboratory. They can link and analyse datasets, while final output is checked against confidentiality by SURS. Usage of the remote access facility costs EUR 18 per month. Researchers can get a limited number of social surveys with less than 1% population included on CDs. In this case data are primary and secondary protected and free of charge.

To get accreditation for using micro data, researchers have to apply for access to micro data, justify the request for micro data and sign a contract and a confidentiality declaration.

In accreditation procedures in SURS different players are involved: the Confidentiality Committee checks the request against the rules for granting access to micro data which are set up and published on SURS's web portal and prepares the proposal for SURS's board of directors. The final decision is taken by the Director-General. It should be mentioned that the Confidentiality Committee works on a weekly basis, so that the service to researchers is really very prompt.

3. Micro data access at the ADP

The ADP offers access to data that are interesting for social science analysis, with emphasis on problems related to the Slovenian society (see more on <http://www.adp.fdv.uni-lj.si/eng/>).

Priority is given to theoretically significant and methodologically well designed studies, especially data gathered over a period of time and international comparative data that include Slovenia. Target users are national and international researchers, teachers and students who are data and statistically literate. General purposes of use are for scientific reports and student seminars. Data are often used as background materials for teaching and practical exercises with students. The ADP also offers data related materials, which are useful for designing new studies (e.g. the original questionnaires and measurement instruments), or can assist in conceptual problem formulation for the secondary data analysis (e.g. links to original publications, based in corresponding data source are available). Users can search and browse through the contents of study descriptions and variables.

4. Cooperation SURS-ADP

The cooperation with the ADP started in 2002 and was enhanced in 2011 due to partnership in the DwB project. Both institutions recognised the synergies in working in the partnership and contributing specialised knowledge. The purpose of cooperation has been recognised in many common activities. The main are to better serve the research community in its need for micro data, to promote and increase micro data use for research purposes, the best contribution to the services in their fields of the specialization and better contribution to the success of the DwB project.

From SURS's perspective the ADP has additional human resources for micro data access service, good knowledge of the DDI metadata standard, additional knowledge about students' needs, more experience and resources for training of researchers and established additional channels for promoting the use of micro data.

Fields of cooperation have been evolving through time and are in fact a division of work with the common aim to increase the use of micro data: preparation of micro data files for use in SURS's on-site laboratory and remote access facility, preparation of non-confidential micro data for less demanding users in ADP's facility, introduction of new methods for data protection particularly suitable for students, preparation of metadata according to the DDI standard with extraction from SURS's metadata systems, communication of metadata according to the DDI standard from both partners, cooperation in user training, promotion of micro data use on SURS and ADP websites, conferences, etc.

5. The results of cooperation with the ADP

From 2003 on, the ADP has been disseminating non-confidential SURS micro data and structured metadata from the Labour Force Survey (LFS), the Household Budget Survey, the Time Use Survey, the Crime Victim Survey and the 2002 Population Census.

The first result of the by DwB enhanced cooperation started in 2011 was LFS micro data set for the last 11 years with the metadata structured according to the DDI standard and prepared for optimised use for researchers with different software packages. The SPSS was used for the exercise, missing values were added to the dataset; additional logical control was made, less important variables were deleted and variables were connected to the codebook. The dataset is promoted on websites of both institutions. Experiences gained in this project will be used in the entire process of micro data and metadata preparation. New data disclosure methods more suitable for inexperienced users were developed too and training of the users is planned for the near future.

6. DwB - towards the common infrastructure for micro data access

The need for global infrastructure was recognised by the EU 7th research program under which the project Data without Boundaries – a four-year EU-funded FP7 project (2011-2015) in which 29 partners are participating – has been launched. Partners are European social science data archives (1/3); NSOs (1/3); universities and other research institutions (1/3) (see more on <http://www.dwbproject.org/>).

The project recognised several obstacles for micro data access such as different frameworks, institutional arrangements, criteria for accreditation, providers, modes of access (no access, safe centres, remote execution, remote access), languages, views about security, anonymisation and output checking procedures.

The aim of the project is linking the capacity of the research community with the important resources of the official statistics micro data in Europe to enhance researchers' access to official statistics micro data in Europe. The focus is on confidential (highly detailed) data and on crossing national boundaries. As a result, new conditions for micro data access at *national level* (strong cooperation between data archives and NSOs, changes in legal framework, providing access to foreign researchers) and at *European level* (central point of access) are expected.

Main issues to be addressed are: building in a sustainable manner the central point of access to official statistics micro data, metadata standards and interoperability, legal issues and accreditation, provision of tools and training researchers for using official statistics micro data.

7. Challenges for the future cooperation at national and international levels

Cooperation between NSOs, data archives and the research community should continue in everyday work. The Slovenian case is a proof that synergies can be achieved in terms of resources but most importantly in terms of better services for the research community. With proper long-term strategies and day-to-day cooperation barriers can be removed. Good cooperation of SURS with the ADP and the research community helped a lot to add the necessary resources for common benefits.

New information technology brings tools which are able to automate the data disclosure procedures. This will save important resources in terms of time and human efforts. Innovative solutions such as the Australian CBS presented by Gare, M. and Chien, C, 2013, are very promising.

The growing need and use of micro data and their trans-border access is a fact and NSOs leaderships should make a strong commitment towards removing all obstacles for providing better access to micro data for research purposes. The DwB project is a good opportunity to make a common effort of data archives, universities, researchers and NSOs towards more research on existing data sets.

SURS believes that in a competitive information environment, NSOs should have the proper strategy and timely action if they want to stay relevant providers of micro data services.

Moreover, cuts of financial and human resources are a reality for a growing number of statistical organisations. As micro data play an important role in the industrialization of statistical production, they are one of the basic elements of an efficient business process. Reuse of micro data internally and externally is one of the pillars for greater efficiency. From this point of view, a

number of NSOs are following the vision of the High Level Group on the Modernisation of Statistical Production and Services (HLG-BAS 2010).

As mentioned, an important element of gaining efficiency is cooperation, sharing resources and data with data archives. Following these strategies will help statistical organisations to strengthen their position in competition with other data producers.

References

CES (2013)“CES Seminar on micro data access”, ECE/CES,
<http://www.unece.org/fileadmin/DAM/stats/documents/ece/ces/2013/>

Gare, M. and Chien, C, 2013 ”Innovative metadata access- Confidentiality on fly”, paper for CES 2013, <http://www.unece.org/fileadmin/DAM/stats/documents/ece/ces/2013/>

HLG-BAS(2010), “Modernisation in official statistics”
<http://www1.unece.org/stat/platform/display/hlgbas/High-level+Group+for+the+Modernisation+of+Statistical+Production+and+Services;jsessionid=6490700DA910EE0AA908CD36651AA946>