Setting up Acquisition Policies for a New Data Archive

*by Helena Laaksonen, Sami Borg & Janez Stebe**

Introduction

Newly established or emerging European social science data archives, like the Finnish Social Science Data Archive (FSD) and the Slovene Social Science Data Archives (ADP), work in national settings where research with secondary data is not deeply rooted. In such a national setting a new service provider must have an active role to prove its value. On the one hand, the archive should relatively quickly supply sufficiently interesting data repository for the scientific community to attract secondary research. On the other hand, it should also be active on the 'demand side' by promoting the usefulness of secondary data in research and training. By raising general awareness about the benefits of secondary research and placing more emphasis on national institutional support and regulation, a culture of data sharing arises in a national setting. This data sharing creates an environment for higher quality acquisition.

Our paper discusses FSD's and ADP's data acquisition policies in relation to their general archival development strategy. We will attempt to systemize the relevance of policy issues by taking a look at current acquisition practices. We will also share experiences from the very grass root level acquisition practices and try to shed light on who are the actors and how their interests play a role in the acquisition. Additionally, we discuss the national institutional support for data sharing. We report the practices of national research funding organizations and present our colleagues' views on how the current practices are operating.

Data archives are typical service sector establishments. A quality of their service can be assessed from users' perspective. How relevant is a service provided by archives to users requirements in type and content. A typical user is a researcher planning secondary analysis; therefore generally the acquisition procedure should be focused on assessing the usefulness of the data for scientific inquiry[1]. National statistics offices have established a set of criteria for the quality of the statistical data serving users requirements that can also be used in the context of a data archive supply. The criteria are accuracy, timeliness, accessibility, comparability, coherence, and finally, cost and burden associated with the products (EUROSTAT 2000). Accuracy means that the data needs to meet certain methodological standards; timeliness that the period between collection and distribution should be short. Accessibility to the data is increased by additional processing and distribution channels used by data archives. The added costs and obligations from the archiving process, and to the data providers, are nevertheless minor in comparison to the whole cost of data collection. The provision of micro data also requires additional confidentiality measures to be established to minimize disclosure risks.

Acquisition and data sharing

Appropriate data acquisition policies support the principles of data sharing and open access to research data. These principles have also recently been highlighted in the OECD declaration of open access to research data from public funding[2]. Widely adopted and efficiently organized data sharing is expected to provide, among other things, better research and science ethics, better quality of research and learning, and more efficient use of public funding. The benefits just mentioned are well argued and accepted in relevant literature. (Hyman 1972; Miller 1977; Royal Statistical Society & the UK Data Archive 2002). However the level of the culture of data sharing is uneven among countries.

Acquisition policies and archival development of a single data archive must be adjusted to its mandate, mission, the resources of the archive, and to its national operational environment. In Finland FSD has had quite a strong position, as it is the first and so far the only national unit whose main aim is to preserve and disseminate research data for secondary use. FSD also has a clear, and too restrictive, mission to archive social science data, and to serve the universities as a national service provider.

As the infrastructure for research with secondary data has emerged at the national level only recently, the national culture for data sharing is still rather weak in most science disciplines and organizations. For this reason, our article also takes a look at the national environment and institutional support for data acquisition.

The role of Slovene ADP in its national setting is similar to the extent that it declares its position as a national social science data archive. This implies certain obligations about the acquisition policy that is implemented in practice. Users would expect to find all important studies in a collection. Whether or not this is the case, it should be accompanied with the appropriate measures to achieve that goal. After attaining sufficient coverage of narrower fields of social science disciplines (sociology, political science and media studies) in its first years of existence, a need may arise to go beyond those disciplines into some of the areas that rely heavily on the empirical data: economy, education, psychology. One can easily locate the main data producing institutions and individuals in those fields, but to establish contacts and share the same understanding of a data sharing culture is yet another story.

Common to both the FSD and the ADP is a rising awareness that a more systematic acquisition policy is needed after the starting period of abundance of data supply, when a whole set of legacy studies were at the priority of acquisition. Now is the time for a systematic evaluation of gaps in a collection and of a more active approach to identify and attract new and important contemporary data sets. After the initial period, a natural claim arises to extend the thematic coverage to disciplines and areas that are partially beyond the limits set at the institutional establishment.

Developments in scope and coverage

Until now FSD's focus has primarily been on archiving quantitative social science data. The present collection includes about 600 datasets, all fully documented with the DDI format. For most datasets, documentation is also available in English[3]. Documentation and processing levels for different types of data were set recently. A new classification and its implementation will hopefully help to adjust future data acquisition policies with the general strategic development of the archive.

Most typical datasets in FSD's collection are nationally representative quantitative surveys from barometers of some non-university national organization. Their time coverage is usually 10 - 15 years. Now the archive is also trying to acquire more data from educational sciences and, as with the ADP, from health sciences, too.

Acquisition of qualitative data started some years ago, but their share of the collection is not more than five per cent. Compared to quantitative surveys, acquisition of qualitative data has proven to be more time consuming and labor consuming. More often than not, the sufficient legal requirements for archiving are not met, as the primary researcher has promised to the persons interviewed that (s)he is the only one who will use the data. Many collectors of qualitative data also tend to think that their data are their personal property. Therefore, in this area, FSD presently aims to affect the overall development of research guidelines and agreements, in order to move them in a direction that would allow easier access to qualitative data.

The depositors in ADP, while mostly academic, do include public and private institutions which gather data themselves. More than 400 studies are currently in the collection. Over half of them are processed to a highest level that includes complete variable level documentation in an electronic Codebook. The study level descriptions are available in English[4].

Initiated with the special project, supported by the Ministry of Information Society in 2002, the ADP started collaboration with the Statistical Office of the Republic of Slovenia and some commercial marketing institutes engaged in data collection. That experience showed that when there is an explicit agreement to share in efforts to process data sets, and when there is additional financial support that covers part of those efforts, then there is also willingness to contribute.

Qualitative data are only sporadic in the Slovene ADP. The archive has not developed special procedures for processing qualitative material. Rather, appropriate sections in DDI format have been used to describe the material and to store the actual qualitative evidence in an electronic format (e.g. scanned paper versions). It is one of the fields that calls for additional efforts in acquisition, as this is the type of data that is not readily available due to different research tradition and corresponding notions of the usefulness of secondary research.

Criteria for evaluating datasets and deciding what to archive

Basically, the FSD and the ADP criteria for evaluation of data is similar to many other European social science data archives. Key evaluation criteria are linked to the scope, to sufficient legal conditions, and to the re-use potential of the data in research and teaching. The ADP was not using strict criteria for selection until 2005 when it introduced a new 'inquiry form' to gain access to rough information about studies that is later on used for selection. Reuse potential of a study is difficult to asses. One may observe current use patterns, where comparative and longitudinal national studies predominate. Experts from a discipline may assist in selection to evaluate scientific potential in collaboration with the archivists, who in turn may make decisions based on administrative criteria. One may also take notice of users suggestions and evaluations of potential studies to be archived.

Data depositors can set different type of conditions for the re-use of their data. Fortunately, this seldomly happens. In the ADP, few of the data sets are under embargo or available only by special permission from the primary researchers. In most cases, the embargo or special

conditions apply only for a certain period of time or to users from nonacademic provenance.

Unlike countries that truly support the idea of open access to research data from public funding, in Finland depositing data or offering it to a data archive is not mandatory in any circumstances. Similarly, in Slovenia collaboration in the whole process from identification to acquisition is entirely voluntary. That means that special effort needs to be invested into the process to negotiate and gain access to the data sets.

On the other hand, the datasets deposited and archived are easily and equally available for researchers and university students. For the moment, the basic data services of the FSD are free of charge for all data providers and end-users. Generally, the FSD does not pay for the data they acquire. Only in very few cases the FSD has paid the costs of preparing data for archive.

The ADP does not charge for educational purposes nor does it charge researchers who don't have institutional support. The charges are minimal for other academic or public use purposes, and full for commercial purposes. In exchange for the data provision, the ADP offers the data providers free access to their data and to an equivalent amount of data provided by other depositors. They require collaboration in the processing stage, but the archive personnel tries to do most of the editing work themselves and asks only for a proofreading of the final version of documentation. These measures, together with a set of guidelines that describe the goals and procedure of data acquisition for potential depositors, are set up to increase willingness to supply data to a data archive by reducing their burden.

Thinking of different types of coverage issues, emphasis has mainly been on data with sufficient, and at least national geographical coverage. The FSD has tried to acquire both panel data and time series that allow examination of trends with several cross sections. Unfortunately, the supply of panel data has been very moderate.

Until now, the work load from updating data and metadata (that has been processed once already) has not yet increased. This allows both of the archives to focus on efforts that aim at a rapid increase of datasets in its holdings. The data evaluation has not yet been very selective if a potential dataset has met the formal requirements posed by the archive.

The ADP aims primarily at collecting theoretically or practically important studies. Studies that fill a research gap or have many implications for a wide range of practical research problems, and have long term scholarly value. Comparative or continuous research with methodological excellence have top priority. In practice, the archive does not reject the data offered, based on the selection criteria. In cases of occasional studies of low quality, the data and materials are just stored as received, without further processing. In active searching for new studies the ADP strives to attain most relevant data, to build high quality collection. The studies that comply with the above mentioned criteria receive the most intensive processing and full electronic documentation. It involves conversion of paper documentation to electronic format and variable level documentation that includes full question texts.

From localisation to receiving the data for archiving

Data acquisition is very labour intensive. Researchers seldom contact the archive expressing a desire to deposit their data for wider use. The staff has to be active and persistent. This is probably the case for most data archives in countries where archiving data is voluntary, and maybe even if it is compulsory. In the following, we share some experiences in acquisition work in Finland and Slovenia.

The acquisition process can be divided into three distinctive stages: identifying potential datasets (we call this stage 'localisation'), negotiating with data creators, and receiving the data and other relevant material. After the last stage, additional information and material are often required to be able to create the metadata and a dataset version suitable for secondary use.

To keep track of all the contacts made regarding any particular dataset, an efficient operative database is needed. Everyone involved in the acquisition process has to enter information on contacts made into the database. Nobody can remember everything, and it may take years from the first contact to the moment when the archive actually receives the material. There are now over a thousand potential datasets recorded in the database of the FSD.

In FSD's new internal operational database[5], the archiving process can be followed right from the identification of a potential dataset to the publication of metadata in the archive's online catalogue.

According to the exchange theory one is most likely to respond to an inquiry 'when perceived costs of doing so are minimised, the rewards are maximised and the expected rewards will be delivered' (Dillman 1983). With the more active acquisition policy in the ADP in 2005 an 'inquiry form' was set. It is suitable for online collection of evidence about potential new studies, so as to 'minimise costs' of providing initial information[6]. A request for giving information was circulated widely to the general user community and specifically to potential data providers. The 'rewards' were emphasised by stressing that providing information to a study's inventory one will contribute to a user-centred data collection with the more exhaustive coverage of high range studies. A modest initial response to this inquiry shows that only additional, more personalised

communication, in most cases e-mails followed by telephone calls, produce positive responses.

The ADP aims to get 30 new studies into the archive each year. In the FSD the goal has been around 80–100 studies, including the international surveys. The ADP, thanks to the inquiry and personal contacts has been able to fulfil the goal in 2005. The studies that are being processed following this initiative will add to topical variety. Almost all of the potential new studies were of highest quality so that no selection was needed except based on the criteria of availability of materials for further processing and distribution. The Finnish archive has reached or almost reached the quite ambitious goal in most years.

Identifying potential contemporary datasets (collected in the 1990s or later)

When identifying potential acquisitions, the FSD is looking for data with potential for secondary use. Primary sources of information are academic journals and the news media. Joining a wide range of e-mail lists has also proven useful, likewise regular browsing through universities' online publication catalogues and the web sites of research funding bodies. The majority of relevant new data can be traced to these sources. Academic literature is less useful, as information published in monographs or in articles of edited collections has usually been published somewhere else previously.

In Slovenia, main sources of information to identify potential new studies are personal contacts with colleagues that are experts in a field and recommend their own or their close collaborators' studies. Other productive sources for first information are news media, as it is widely adopted practice to have a news conference after the first results of a research project are published. Other sources are SICRIS (Slovene Current Research Information System)[7] that covers publicly funded research projects since 1998. Updates on information gathered through the course of the research project are often missing, including if a project is of empirical nature and if a data set exists at all. The other information system is COBISS[8] which is a national record-keeper of all Slovenian libraries and includes an up-to-date bibliography of registered researchers. Institutional home pages are another source of information, useful in particular when one knows what one is looking for and to add reference to initial information, gained from some other sources. The ADP also keeps and updates information on availability of some international data sets kept in other national archives or in specialised ad hoc project archives. All in all, work on data acquisition is labour intensive and demanding task already in a stage of identification of datasets.

Detective work to preserve older data

When it comes to tracking down older research material, literature reviews, and particularly the contacts with experienced academics, are important. One can naturally browse through older issues of academic journals to acquire information on data colleted during past decades. However, without personal contacts it is often difficult to discover the present whereabouts of an older dataset.

In Finland, the data collected prior to the 1990s, even when identified, very seldom end up at the archive. It may be laborious to migrate the data into a present day format. Still, this is mostly manageable. What often makes the archiving of older data impossible is that available metadata is not at all adequate. If the archive cannot find out what the variables mean, who the respondents were and how the sample was drawn, the data cannot be used.

In Slovenia there were some cases when no data set was available any more for some past studies, and the archive tried to preserve at least what was remaining on paper, such as reports and questionnaires.

As it has proven difficult to acquire older data, its share in the collections is a minor one. Around a half of the archived quantitative data were collected in the 1990s, in both of the archives, and around one third since the year 2000. Thus, getting data collected in the 2000s has not been too difficult.

Contacting potential depositors

At the FSD, all personnel are involved in localisation. The main responsibility for getting data into the archive, however, lies on the shoulders of the Director and the Information Officer. They are the ones who usually contact the data creators. The Research Officer in FSD is a specialist in research ethics and qualitative research. She is mainly responsible for acquisition of qualitative data.

At the ADP, the personnel consists of only two persons. The initial contacting of the depositors is mainly the Director's duty, but further acquiring and processing of new studies is done by both.

Acquiring data is a tough job

Convincing researchers takes a lot of persuasion, whether in Finland or Slovenia. One has to repeat the same reasoning over and over again, to the same person or maybe to several persons, as it sometimes takes time to find the right person. There are two main stages. First, one needs to reach the 'ok, I will give you the data' agreement stage. This does not necessarily get the process much further. Encouraging the researcher to act and actually transfer the data into the archive is another story.

At the FSD, in one relatively easy negotiation process, where the dataset was deposited within a year of the first contact, altogether 10 contacts were made (e-mails and phone calls). In some cases it has taken only about half a year from the first contact to the arrival of the data. Still,

several e-mails have to be sent and many phone calls made, before the actual delivery of the data. In most cases, researchers agreed to archive their data at the first contact. It was much more difficult to get them to actually send the material and deposit agreements to the archive.

The biggest negotiation challenge is to persuade an individual researcher to archive a one-time study. If the FSD already has established contact with someone in an organisation regularly collecting data, the job may be easier but not necessarily. Established contacts must be maintained. Generally speaking, it is easier to deal with medium-sized organisations than with large ones. Why? A probable reason is that large organisations lack the culture of archiving and data sharing. It is also difficult to find persons in large organisations who would take responsibility for these matters.

Recently collected datasets, with primary analyses made and published, are easier to acquire than older research material, which may have been put aside somewhere. However, there are a couple of moments when the archive has good chances of success: when researchers are approaching retirement age or when they move office. At these times researchers may be willing to let their beloved research data fall into the hands of the archive staff.

At the ADP experiences are similar to the extent that it is hard to persuade researchers to move forward. After the first contact and when usual arguments are being exchanged about the purpose and benefits of having the data set offered to the data archive, in most cases researchers are willing to provide their data. What is a main obstacle in that even finding the data files and collecting the documentation may take additional time that is extremely scarce. That adds to an argument that some form of institutional legal obligation accompanied with the financial support would probably help to reduce 'perceived costs'. In addition, the proper reference to an authority increases mutual trust that is a basis for expecting that 'rewards will be delivered'.

PR and information services

An essential part of acquisition, at least for fairly recently founded archives like the ADP and the FSD, is to make the archive known, and to promote new ways of thinking within the research community. Researchers have all kinds of reasons why they do not want to archive their data for re-use. Everyone with experience of data acquisition is familiar with some of the reasons. Qualitatively oriented researchers are the toughest cases. They need to be addressed with particular care to make them see data sharing as an essential part of the research process.

In making the archive known, standard PR methods are used. However, approaching the social science research community may be tricky. One has to rouse their interest, but not appear too 'commercial'. One needs to provide information, but as an expert rather than as a salesperson. This is not necessarily easy to carry out.

The FSD distributes general information on its services through its own channels, and the personnel write short articles for the publications of other organisations and associations. Newsletters on recently published data and other services are sent through FSD's e-mail list and by post to selected target persons and groups. New generations of students and future researchers are also an important target group, as are university teachers. Contacts with key persons, university libraries and research institutes are important. These bodies are informed of the data archive's activities and collections through all possible channels. In addition, the archive staff visit them, invite them to visit the archive or to attend the conferences and seminars organised by the FSD.

Conference presentations and posters, articles in scientific journals, lectures for students of different faculties, and other channels of information about the ADP's activities and holdings are more often being conceived as an essential part of its activity. That this is not an easy task shows evidence that in personal contacts the ADP personnel still encounter researchers who are unfamiliar with the mission of the ADP. Most often, the reaction after explaining the basic principles of data preservation and reuse, is exclamation of surprise and of support. The ADP plans to integrate into its activities thematically dedicated workshops. These workshops will share the specialised practical knowledge of the eminent researchers, who have long term experiences with the analysis of particular data sets, with the wider community of users. In the end, only intensive contacts and personalised communication have been shown to achieve the desired results.

To promote a new culture of data sharing within the research community, it is also necessary to proclaim the virtues of open access, transparency, possibility of replication and validation of research, and other good research practices. These issues must be brought up repeatedly when contacting individual researchers.

However, caution is needed to not get the opposite result from the one intended. As mentioned, researchers are a tricky bunch of people. It is often useful to appeal to the researchers' own interests. What are the advantages for them? A free-of-charge, reliable preservation for their data, and of course, fame and glory, when other people get acquainted with their excellent data, and cite them. However, using the latter argument may sometimes backfire. Once the principal researcher told to an FSD employee that the data was 'so bad', i.e., of so low quality that he did not want to share it – and refused to archive it on those grounds. On that ground one may claim that willingness to provide the data to a data archive is also a guaranty of its consistency and overall quality, that can

be further tested by secondary analysts. Which is another argument that public funding agencies could use to make the data sharing a legal obligation. After all, the data are made on public money and with the collaboration of lay people.

Support from the funding organisations in Europe
During the spring of 2005 the FSD turned to 20 European data archives to collect experiences of what kind of support social science funding organisations give to archiving and data sharing. Nine data archives sent their answers. A short summary and some examples of practice are provided below. The FSD thanks all contributors[9].

We asked: 1) What kind of guidelines or regulations the research funding organisations have on the archiving, accessibility and re-use of research data, 2) in what kind of a document do the guidelines appear, and 3) are the guidelines implemented, i.e. are researchers doing what the guidelines say they should do.

In all but one country (Italy) at minimum one funding body has at least a recommendation to deposit data for archiving. In some cases it was difficult to interpret whether it was a recommendation or a requirement. According to our interpretation, in five countries the funding body or bodies had taken a stronger stance for archiving than mere recommendations.

However, even when the funding is given 'on the condition that data be deposited' there are seldomly sufficient effective ways of controlling that the demand is fulfilled. The most effective means, a sanction actually, is set in the UK, where the project might not receive part of its funding, if it has not offered its data to the data archive. This applies to funding received from The Economic and Social Research Council (ESRC). Accordingly, the guidelines are implemented at a much higher level than earlier.

Recommendations or not, very often it seems to be the archives who have to do the policing afterwards. Yet, in some countries the funding body itself checks to determine whether or not its policy is followed by the research projects. For example, in Germany, National Science Foundation (DFG) monitors final project reports to check whether data has been deposited and the Science council checks in its evaluations of institutes, whether they conform to policy recommendations. In Switzerland, the Swiss National Science Foundation (SNSF) selects a subset of projects in which they explicitly ask the Primary Investigator to take contact with SIDOS to discuss the opportunity of depositing the data for further use and to include information on that contact in the first intermediate report. If the report is missing, a reminder is sent to the researcher. The SIDOS evaluates the case together with the PI and makes a recommendation. The data on which SIDOS and PI agree are usually deposited.

We were also interested in financial support for research projects to prepare data for archiving and re-use. We asked whether researchers include expenses for preparing data for archiving and re-use in their research proposals to funding bodies. We also asked whether they would get the money if they applied for it.

There is only one country where the research projects usually apply for money to prepare data for archiving. That is the UK, where they also have a sanction for not depositing the data. In five countries, researchers almost never or never apply funds for this purpose. In one country they sometimes do apply, and in one they seldom do. Yet, seven of our respondents estimated that, if applied for, funding would usually be given to the projects. In three countries (according to our interpretation of the answers) they would usually not get the money.

Further, we asked whether the archives have tried to influence the research funding organisations' policies regarding data sharing and archiving, and regarding financial support for preparing data for archiving. In all countries, the archives have been and will have to continue to be more or less active in lobbying for the cause of archiving and data sharing.

It is clear that newer archives have a longer way to go to receive full institutional recognition and support from national legislation. The role that the data archives have is similar to well established 'national heritage' institutions like national libraries, museums and classical archives. New data archives could make a step further in this direction with the help of examples of good practice in countries where infrastructural role of a data archive is to support high quality scientific production and high quality education.

Concluding remarks
At present, both the FSD and the ADP are clearly moving from the first phase of a newly established archive to a second phase. At first, they concentrated on building up the data collection, and could not afford to be very selective. They needed to be active, and contacted a large number of persons and organisations within respective research communities.

Now the reserve of potential, unidentified datasets is diminishing within the core social science disciplines. But at the same time the data preservation needs of other disciplines are growing. Therefore the FSD will, based on its own decisions, expand its coverage to some related research fields, mainly health and educational sciences. Including more qualitative data into the holdings clearly is also a very important choice of acquisition during this second phase of development. The ADP intends to cover the fields mentioned and in addition intends to explore more intensive collaboration with economists and

psychologists.

The archives hope to see the third phase of development soon. That would require more institutional support for them, especially from research funding organisations. To ensure a steady flow of data into the archive, new institutional policies are needed. The projects and current policies aim at establishing a research culture where data sharing is considered an inherent part of a research project – to be taken into account already when preparing research proposals. The cost of preparing data for archiving should be included in research applications. Researchers are to be advised to include into contracts and communication with sponsors, and human subjects who collaborate, an explicit agreement about data delivery to the data archive for the purpose of secondary analysis.

Both online and printed material will be produced to promote this view, and all these actions will be linked to the national implementation of the OECD guidelines (OECD 2004). Major research funding bodies should be encouraged to place more weight on the issue, and to include in their funding decisions recommendations or requirements for the data to be archived.

When implemented, such institutional changes would improve the overall efficiency of acquisition in a strategy, which can set the requirements of a quality and relevance for new studies higher. That can be accomplished if the coverage of the studies accessible for acquisition would be almost complete, that is, if information about studies would be supplied regularly and data ready for processing without further obstacles. Researchers already in a planning stage of a project would be advised to include an option for giving data to an archive when negotiating a contract with sponsors, to include initial preparation of data and documentation among the tasks, and communicate with the research subjects about the secondary use.

If all this will happen in the near future, our next paper on data acquisition might concentrate more on the issue of designing and improving the collection. The portion of staff time that is currently devoted to initial negotiations about sharing the data could be concentrated on quality of processing and distribution instead.

References

Dillman, Don A. (1983): Mail and Other Self-Administred Questionnaires." In: Rossi, P.H., J.D. Wright, A.B. Anderson (Eds.): Handbook of Survey Research. Academic Press, Inc., San Diego. 359-378.

EUROSTAT (2000): Standard Quality Report. Assessment of the Quality in Statistics. Luxemburg, 4-5/04/2000. www.unece.org/stats/documents/2000/11/metis/crp.3.e.pdf

Gutmann, M., K. Schürer, D. Donakowski and H. Beedham (2004): The selection, appraisal, and retention of social science data. Data Science Journal, 3, 2004. http://journals.eecs.qub.ac.uk/codata/Journal/contents/3_04/3_04pdfs/DS386.pdf

Hyman, H.H. (1972): Secondary Analysis of Sample Surveys. New York: John Wiley & Sons, Inc.

Miller, W.E. (1977): "The Less Obvious Functions of Archiving Survey Research Data". In: Hofferbert, R.I., J.M. Clubb (Eds.) (1977): Social Science Data Archives. Beverly Hills: Sage.

Mochmann, E., P. de Guchteneire (1988): "Data services for the Social Sciences". In: L. Kiuzadjan, K.T. Saelen, G. Soloviev: Information needs, problems and possibilities. Vienna: European Coordination Centre for Research and Documentation in the Social Sciences (Vienna Centre). http://www.ifdo.org/data/data_archive_workflow01.html

OECD (2004): Science, Technology and Innovation for the 21st Century. Meeting of the OECD Committee for Scientific and Technological Policy at Ministerial Level, 29-30 January 2004 - Final Communique. http://www.oecd.org/document/0,2340,en_2649_34487_25998799_1_1_1_1,00.html

Royal Statistical Society & the UK Data Archive (2002) Preserving & Sharing Statistical Material. Working Group on the Preservation and Sharing of Statistical Material: Information for Data Producers. Colchester: UK Data Archive, University of Essex.

Sivonen, Jouni: (2004) New User Interface for Managing the Archiving Process in FSD. Paper presented at IASSIST Conference, Madison, May 2004. See: http://www.iassistdata.org/conferences/2004/presentations/C3_sivonen.ppt

* The article is based on two presentations at the IASSIST conference, Edinburgh, 25 May, 2005 in Session A3: Enlightened Policies: Improving Collections and Acquisitions. The presentations were pulled together and updated in 2006. Helena Laaksonen & Sami Borg, Finnish Social Science Data Archive, Janez Stebe, Social Science Data Archives, University of Ljubljana, Slovenia. email: helena.laaksonen@uta.fi, sami.borg@uta.fi, janez.stebe@fdv.uni-lj.si

Endnotes

1 "A general guideline is whether on not the data are usable for future scientific research" (Mochman and Guchteneire 1988); "The extent to which the data will advance knowledge" (Guttman et al. 2004).

2 OECD 2004. After the declaration the member countries have set up a working group to specify possibilities and

practices for the implementation.

3 See more at http://www.fsd.uta.fi/english/data/

4 See http://www.adp.fdv.uni-lj.si/opisi/

5 FSD's operational database was presented at IASSIST 2004 (Sivonen).

6 See http://www.adp.fdv.uni-lj.si/edan/.

7 http://sicris.izum.si/

8 http://cobiss.izum.si/

9 The inquiry was sent to 20 data archives in Europe, and nine responded. The answers were provided by Ekkehard Mochmann for ZA, Germany; Reto Hadorn for SIDOS, Switzerland; Susan Cadogan for the UK Data Archive; Hans Jørgen Marker for DDA, Denmark; Janez Stebe for ADP, Slovenia; Carlo Pisano for ADPSS-Sociodata, Italy; Iris Alfredsson for SSD, Sweden; Marion Wittenberg for Steinmetz Archive, the Netherlands; and Gry Henriksen for NSD, Norway. A more comprehensive overview of the answers was given in the APPENDIX to the actual IASSIST paper by S. Borg & H. Laaksonen (2005). It can be attained from the FSD.